U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

PB80-131485

Source Assessment: Analysis of Uncertainty - Principles and Applications

Monsanto Research Corp, Dayton, OH

Prepared for

Industrial Environmental Research Lab, Research Triangle Park, NC

Aug 78

029195

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-600/2-78-004u
2.
3. RECIPIENT'S ACCESSION NO. PB 80 131485
4. TITLE AND SUBTITLE SOURCE ASSESSMENT: Analysis of Uncertainty--Principles and Applications
5. REPORT DATE August 1978
6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) R. W. Serth, T. W. Hughes, R. E. Opferkuch, and E. C. Eimutis
8. PERFORMING ORGANIZATION REPORT NO. MRC-DA-632
9. PERFORMING ORGANIZATION NAME AND ADDRESS

Monsanto Research Corporation
1515 Nicholas Road
Dayton, Ohio 45407

10. PROGRAM ELEMENT NO. 1AB015; ROAP 21AXM-071
11. CONTRACT/GRANT NO. 68-02-1874
12. SPONSORING AGENCY NAME AND ADDRESS

EPA, Office of Research and Development
Industrial Environmental Research Laboratory
Research Triangle Park, NC 27711

13. TYPE OF REPORT AND PERIOD COVERED Task Final; 11/76-3/78
14. SPONSORING AGENCY CODE EPA/600/13
15. SUPPLEMENTARY NOTES IERL-RTP project officer is Dale A. Denny, Mail Drop 62, 919/541-2547. Similar previous reports are in the EPA-600/2-76-032 and -77-107 series.

16. ABSTRACT The report gives results of: an analysis of the uncertainties involved in calculating the decision parameters used in the Source Assessment Program; and a determination of the effect of these uncertainties on the decision-making procedure. A general procedure for performing an analysis of uncertainty is developed, based on the principles of error propagation and statistical inference. It is shown that this simple and straightforward method represents and approximation to standard statistical techniques. The approximate method is illustrated by application to four problems in the area of environmental control. The general procedure is used to establish guidelines for conducting air emissions studies in the Source Assessment program. In particular, guidelines are established for precision in field sampling and analytical work, and for setting critical values of decision parameters.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS		b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Pollution	Sampling	Pollution Control	13B
Uncertainty Principle	Analyzing	Stationary Sources	20J
Statistical Inference		Source Assessment	12A
Statistical Distributions		Error Propagation	
Decision Making			05A
Field Tests			14B

18. DISTRIBUTION STATEMENT Unlimited
19. SECURITY CLASS *(This Report)* Unclassified
20. SECURITY CLASS *(This page)* Unclassified
21. NO. OF PAGES 181
22. PRICE A09-A01

United States
Environmental Protection
Agency

Industrial Environmental Research
Laboratory
Research Triangle Park NC 27711

EPA-600/2-78-004u
August 1978

PB80-131485

Research and Development

EPA

Source Assessment: Analysis of Uncertainty - Principles and Applications

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U. S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

RESEARCH REPORTING SERIES

Research reports of the Office of Research and Development, U.S. Environmental Protection Agency, have been grouped into nine series. These nine broad categories were established to facilitate further development and application of environmental technology. Elimination of traditional grouping was consciously planned to foster technology transfer and a maximum interface in related fields. The nine series are:

1. Environmental Health Effects Research
2. Environmental Protection Technology
3. Ecological Research
4. Environmental Monitoring
5. Socioeconomic Environmental Studies
6. Scientific and Technical Assessment Reports (STAR)
7. Interagency Energy-Environment Research and Development
8. "Special" Reports
9. Miscellaneous Reports

This report has been assigned to the ENVIRONMENTAL PROTECTION TECHNOLOGY series. This series describes research performed to develop and demonstrate instrumentation, equipment, and methodology to repair or prevent environmental degradation from point and non-point sources of pollution. This work provides the new or improved technology required for the control and treatment of pollution sources to meet environmental quality standards.

EPA REVIEW NOTICE

This report has been reviewed by the U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policy of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

EPA-600/2-78-004u
August 1978

Source Assessment:

Analysis of Uncertainty - Principles and Applications

by

R.W. Serth, T.W. Hughes, R.E. Opferkuch, and E.C. Eimutis

Monsanto Research Corporation
1515 Nicholas Road
Dayton, Ohio 45407

Contract No. 68-02-1874
ROAP No. 21AXM-071
Program Element No. 1AB015

EPA Project Officer: Dale A. Denny

Industrial Environmental Research Laboratory
Office of Energy, Minerals, and Industry
Research Triangle Park, NC 27711

Prepared for

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Research and Development
Washington, DC 20460

PREFACE

The Industrial Environmental Research Laboratory (IERL) of EPA has the responsibility for insuring that pollution control technology is available for stationary sources to meet the requirements of the Clean Air Act, the Federal Water Pollution Control Act and solid waste legislation. If control technology is unavailable, inadequate, or uneconomical, then financial support is provided for the development of the needed control techniques for industrial and extractive process industries. The Chemical Processes Branch of the Industrial Processes Division of IERL has the responsibility for investing tax dollars in programs to develop control technology for a large number of operations (more than 500) in the chemical industries.

Monsanto Research Corporation (MRC) has contracted with EPA to investigate the environmental impact of various industries which represent sources of pollution in accordance with EPA's responsibility as outlined above. Dr. Robert C. Binning serves as MRC Program Manager in this overall program entitled "Source Assessment," which includes the investigation of sources in each of four categories: combustion, organic materials, inorganic materials, and open sources. Dr. Dale A. Denny of the Industrial Processes Division at Research Triangle Park serves as EPA Project Officer. Reports prepared in this program are of three types: Source Assessment Documents, State-of-the-Art Reports and Special Project Reports.

Source Assessment Documents contain data on emissions from specific industries. Such data are gathered from the literature, government agencies and cooperating companies. Sampling and analysis are also performed by the contractor when the available information does not adequately characterize the source emissions. These documents contain the information necessary for IERL to decide whether a reduction in emissions is required.

State-of-the-Art Reports include data on emissions from specific industries which are also gathered from the literature, government agencies and cooperating companies. However, no extensive sampling is conducted by the contractor for such industries. Results from such studies are published as State-of-the-Art Reports for potential utility by the government, industry, and others have specific needs and interests.

Special projects provide specific information or services which are applicable to a number of source types or have special utility to EPA but are not part of a particular source assessment study. This special project report, "Source Assessment: Analysis of Uncertainty, Principles and Applications," was prepared to describe a procedure for analyzing the uncertainty associated with data and other information contained in source assessment studies. The general principles and procedures used in the analysis are illustrated by means of four applications to decision making in the area of environmental control. The application of these principles to the study of air emissions in the Source Assessment Program is described in detail.

ABSTRACT

This report provides the results of a study that was conducted to analyze the uncertainties involved in the calculation of the decision parameters used in the Source Assessment Program and to determine the effect of these uncertainties on the decision-making procedure.

A general procedure for performing an analysis of uncertainty is developed based on the principles of error propagation and statistical inference. It is shown that this simple and straightforward method represents an approximation to standard statistical techniques. The approximate method is illustrated by application to four problems in the area of environmental control.

The general procedure is used to establish guidelines for conducting air emissions studies in the Source Assessment Program. In particular, guidelines are established for precision in field sampling and analytical work, and for setting critical values of decision parameters.

This report was submitted in partial fulfillment of Contract No. 68-02-1874 by Monsanto Research Corporation under the sponsorship of the U.S. Environmental Protection Agency. The study described in this report covers the period November 1976 to March 1978.

CONTENTS

Preface iii
Abstract v
Figures ix
Tables x
Abbreviations and Symbols xi

1. Introduction 1
2. Summary 4
    - Background 4
    - Source severity 4
    - National emission burden 5
    - Offset calculations 5
    - Reasonably available control technology 6
3. Statistical Basis for the Procedure 8
4. Source Assessment Methodology 13
    - Definition of indices 13
    - Expanded discussion of source severity 15
    - Representative source 17
    - Source severity distributions 19
    - Stochastic approach to source severity 19
5. Source Assessment: Source Severity 22
    - Introduction 22
    - Uncertainties in parameters 23
    - Deterministic decision approach 25
    - Stochastic decision approach 28
    - Tests of hypothesis 30
    - Comparison of alternate decision approaches 35
    - Guidelines for source assessment program 35
    - Allowable random uncertainty in emission factor . . . 39
    - Effect of random uncertainty in emission factor . . . 40
    - Operating characteristics of the test 42
    - Summary 46
6. Source Assessment: Emissions Burden 47
    - Introduction 47
    - Uncertainty in N_B 48
    - Statistical test of hypothesis 50
    - Summary 51
7. Offset Calculations: Plant Expansion Problem 52
    - Introduction 52
    - Working equations 52
    - Uncertainty in $(Q_2 - Q_1)$ 53
    - Test of hypothesis 55
    - Numerical Example 56
    - Summary 57

CONTENTS (continued)

8. Comparison of Alternative Controls: Reasonably Available Control Technology 59
Introduction 59
Governing equations 59
Test of hypothesis 61
Numerical example 62
Summary 63
9. Conclusions and Recommendations 64
Conclusions 64
Recommendations 66

References 68

Appendices

A. Accuracy, error, and uncertainty 74
B. Use and interpretation of error propagation formulas . . 77
C. Derivation of source severity equations 98
D. Plume rise 110
E. Alternative methods for estimating "acceptable" concentration for noncriteria pollutants 116
F. Construction of operating characteristic curves 120
G. Relationship of sampling and analysis bias to systematic errors 127
H. Source severity simulation and probabilistic sensitivity analysis 129
I. Quantification of uncertainty in source severity 139
J. Example calculations for source assessment: carbon black manufacture 151

Glossary 154

Conversion Factors and Metric Prefixes 156

FIGURES

Figure		Page
1.	Deterministic source severity distribution for carbon monoxide emissions from carbon black plants	20
2.	Impact of alternative hypothesis test approaches on S_C^*	36
3.	Variation in critical source severity as a function of emission factor data quality and the acceptable number of days for severity to exceed 1.0 when treating the true source severity as a random parameter	36
4.	Schematic representation of the fiducial statistical approach	42
5.	Operating characteristic curves for Case A with $\hat{b}_r = 0.10$	43
6.	Effective operating characteristic curve for Case A with $\hat{b}_r = 0.10$	44
7.	Effective operating characteristic curves for Case A as a function of random uncertainty in emission factor	45
8.	Operating characteristic curve for the stochastic decision test approach	45
9.	Effective operating characteristic curve for example problem	57

TABLES

Number — Page

1. Measure/Analysis Aspects of Source Severity 5
2. Measure/Analysis Aspects of the National Emissions Burden 6
3. Generalized Error Bounds for Source Assessment 37
4. Contribution of Individual Uncertainties to Total Uncertainty in Source Severity (Deterministic Decision Approach) 40
5. Contribution of Individual Uncertainties to Error Bounds on Source Severity (Deterministic Decision Approach) 41

ABBREVIATIONS AND SYMBOLS

SECTIONS 1 THROUGH 5

A -- estimated value of average production capacity, kg/s

a_r -- random uncertainty associated with A, kg/s

$\hat{a}_r$ -- a_r/A

B -- measured value of emission factor, g/kg

b -- bias

b_ℓ, b_u -- bounds on bias

b_r -- random uncertainty associated with B, g/kg

b_{r1} -- uncertainty in emission factor due to imprecision in sampling, g/kg

b_{r2} -- uncertainty in emission factor due to imprecision in analysis, g/kg

b_{r3} -- uncertainty in emission factor due to process variation, g/kg

b_{r4} -- uncertainty in emission factor due to imprecision in measurement of production rate, g/kg

$\hat{b}_r$ -- b_r/B

$\hat{b}_{s\ell}$ -- $b_{s\ell}/B$

b_{su}, $b_{s\ell}$ -- upper and lower systematic uncertainties associated with B, g/kg

$\hat{b}_{su}$ -- b_{su}/B

C -- estimated value of effective emission height, m

CAP -- production capacity, kg/s

CO -- carbon monoxide

c_r - random uncertainty associated with C, m

$\hat{c}_r$ -- c_r/C

$\hat{c}_{s\ell}$ -- $c_{s\ell}/C$

c_{su}, $c_{s\ell}$ -- upper and lower systematic uncertainties associated with C, m

$\hat{c}_{su}$ -- c_{su}/C

D -- estimated value of "acceptable" concentration, g/m^3

ABBREVIATIONS AND SYMBOLS (continued)

$\hat{d}_{s\ell}$ -- $d_{s\ell}/D$
d_{su}, $d_{s\ell}$ -- upper and lower systematic uncertainties associated with D, g/m^3
$\hat{d}_{su}$ -- d_{su}/D
e -- natural base logarithm, $\cong$ 2.72
$E(\hat{\underline{\chi}})$ -- expected value of $\hat{\underline{\chi}}$ (underbar denotes random variable)
EF -- emission factor, g/kg
e_r -- random uncertainty associated with S_C
$\hat{e}_r$ -- e_r/S_C
$\hat{e}_{s\ell}$ -- $e_{s\ell}/S_C$
e_{su}, $e_{s\ell}$ -- upper and lower systematic uncertainties associated with S_C
$\hat{e}_{su}$ -- e_{su}/S_C
$\hat{e}_u$, $\hat{e}_\ell$ -- upper and lower total relative uncertainties associated with S_C
F -- "acceptable" pollutant concentration, g/m^3
f_1, f_2 -- weighting factors
G -- factor which converts TLV to "equivalent" PAAQS
h -- physical stack height, m
H -- effective emission height, m
H_0 -- null hypothesis
H_A -- alternate hypothesis
I_C -- calculated value of decision index
I_T -- true value of decision index
k -- constant, s/m
m -- 97.5% point of distribution of γ
M_n -- annual mass emissions of given criteria pollutant from all stationary sources nationwide, metric tons
M_P -- annual mass emissions of given criteria pollutant from given source type, metric tons
M_s -- annual mass emissions of given criteria pollutant from all stationary sources within a given state, metric tons
n, n_1, n_2 -- sample size
$\bar{n}$ -- fraction of a year
NO_x -- nitrogen oxides
PAAQS -- primary ambient air quality standard, g/m^3

ABBREVIATIONS AND SYMBOLS (continued)

Q -- mass emission rate, g/s
r -- number of days per year
s, s_1, s_2 -- sample standard deviation
S -- source severity
S_C -- "calculated" source severity
S_C^* -- critical test value of calculated source severity
SO_x -- sulfur oxides
S_T -- "true" source severity
t -- averaging time, min; or student t value
t_{ave} -- weighted average Student t value
TLV -- threshold limit value, g/m^3
t_o -- short-term averaging time, 3 min
V(x) -- variance of x
u -- wind speed, m/s
α -- level of test
γ -- correction factor for Gaussian plume equation
ΔH -- plume rise, m
μ, μ_1, μ_2 -- population mean
σ -- population standard deviation
$\hat{\chi}$ -- estimate of population mean
$\bar{\chi}$, $\bar{\chi}_1$, $\bar{\chi}_2$ -- sample mean
$\bar{\chi}_{max}$ -- maximum time-averaged ground level concentration, g/m^3
$\bar{\chi}_{meas}$, $\bar{\chi}_{pred}$ -- measured and predicted values of ground level concentration, g/m^3
Z -- number of days per year

SECTION 6

A -- measured value of CAP_T, kg/yr
$\hat{a}_r$ -- relative random uncertainty associated with A
B -- measured value of EF_R, g/kg
$\hat{b}_r$ -- relative random uncertainty associated with B
b_{su}, $b_{s\ell}$ -- upper and lower systematic uncertainties associated with B, g/kg
CAP_T -- total production capacity of given source type, kg/yr
D -- measured value of M_n, kg
$\hat{d}_r$ -- relative random uncertainty associated with D

ABBREVIATIONS AND SYMBOLS (continued)

d_{su}, $d_{s\ell}$ -- upper and lower systematic uncertainties associated with D, kg

EF_R -- representative emission factor for given source type, g/kg

$\hat{e}_r$ -- relative random uncertainty associated with $(NB)_C$

$\hat{e}_{su}$, $\hat{e}_{s\ell}$ -- upper and lower systematic uncertainties associated with $(NB)_C$

$\hat{e}_u$, $\hat{e}_\ell$ -- relative upper and lower total uncertainty in N_B

H -- null hypothesis

M_n -- annual mass emissions of a given criteria pollutant from all stationary sources nationwide, kg

M_{NEDS} -- estimate of M_n obtained from 1972 National Emissions Report, kg

M_p -- annual mass emissions of a given criteria pollutant from a given source type, kg

N_B -- national emissions burden

$(N_B)_C$ -- calculated value of N_B

$(N_B)_T$ -- true value of N_B

$(N_B^*)_C$ -- critical test value of $(N_B)_C$

α -- level of test

SECTION 7

b_{rn} -- random uncertainty associated with B_N, g/s

b_{rNi} -- random uncertainty in Q_{Ni}, g/s

b_{r1} -- random uncertainty associated with B_1, g/s

$b_{r1'}$ -- random uncertainty associated with B_1', g/s

b_{r1j} -- random uncertainty in Q_{1j}, g/s

$b_{r1'j}$ -- random uncertainty in $Q_{1'j}$, g/s

$b_{r\Delta}$ -- random uncertainty associated with B_Δ, g/s

b_{un}, $b_{\ell N}$ -- upper and lower systematic uncertainties associated with B_N, g/s

b_{uNi}, $b_{\ell Ni}$ -- upper and lower systematic uncertainties associated with Q_{Ni}, g/s

b_{u1}, $b_{\ell 1}$ -- upper and lower systematic uncertainties associated with B_1, g/s

$b_{u1'}$, $b_{\ell 1'}$ -- upper and lower systematic uncertainties associated with B_1', g/s

ABBREVIATIONS AND SYMBOLS (continued)

b_{u1j}, $b_{\ell 1j}$ -- upper and lower systematic uncertainties in Q_{1j}, g/s

$b_{u1'j}$, $b_{\ell 1'j}$ -- upper and lower systematic uncertainties in $Q_{1'j}$, g/s

$b_{u\Delta}$, $b_{\ell\Delta}$ -- upper and lower systematic uncertainties associated with B_Δ, g/s

B_N -- measured value of Q_N, g/s

B_1 -- measured value of Q_1, g/s

$B_{1'}$ -- measured value of $Q_{1'}$, g/s

B_Δ -- measured value of Δ_1

e_r -- random uncertainty associated with difference $B_N - B_\Delta$, g/s

e_{su}, $e_{s\ell}$ -- upper and lower systematic uncertainties associated with $B_N - B_\Delta$, g/s

e_u, e_ℓ -- upper and lower total uncertainties associated with $B_N - B_\Delta$, g/s

H_A -- alternative hypothesis

H_0 -- null hypothesis

Q_N -- emission rate from new unit, g/s

Q_{Ni} -- emission rate from ith emission point in new unit, g/s

Q_1 -- original emission rate, g/s

$Q_{1'}$ -- emission rate from original unit after expansion, g/s

Q_{1j} -- emission rate before expansion from jth emission point in original unit, g/s

$Q_{1'j}$ -- emission rate after expansion from jth emission point in original unit, g/s

Q_2 -- emission rate after expansion, g/s

$(Q_2 - Q_1)_C$ -- calculated value of $Q_2 - Q_1 = B_N - B_\Delta$, g/s

$(Q_2 - Q_1)_T$ -- true value of $Q_2 - Q_1$, g/s

ΔQ_1 -- $Q_2 - Q_{1'}$, = reduction in emission rate from original unit, g/s

$(\Delta Q_1)_j$ -- difference in emission rate before and after expansion from jth emission point in original unit, g/s

α -- level of test

b_{r1} -- random uncertainty in Q_1, g/s

ABBREVIATIONS AND SYMBOLS (continued)

SECTION 8

$\hat{b}_{r1} = b_{r1}/B_1$ etc. -- relative random uncertainties

b_{r2} -- random uncertainty in Q_2, g/s

b_{u1}, $b_{\ell 1}$ -- upper and lower systematic uncertainties in Q_1, g/s

$\hat{b}_{u1} = b_{u1}/B_1$ etc. -- relative systematic uncertainties

b_{u2}, $b_{\ell 2}$ -- upper and lower systematic uncertainties in Q_2, g/s

B_1 -- measured value of Q_1, g/s

B_2 -- measured value of Q_2, g/s

e_{su}, $e_{s\ell}$ -- upper and lower systematic uncertainties in $(\varepsilon_2 - \varepsilon_1)$

e_u, e_ℓ -- upper and lower total uncertainties in $(\varepsilon_2 - \varepsilon_1)$

H_A -- alternative hypothesis

H_0 -- null hypothesis

Q_1 -- uncontrolled emission rate, g/s

Q_2 -- controlled emission rate with presently installed device, g/s

α -- level of test

β_ℓ, β_u -- lower and upper systematic uncertainties in ε_2

β_r -- random uncertainty in ε_2

ε_1 -- control efficiency of presently installed device

ε_2 -- control efficiency of RACT

$(\varepsilon_2)_C$ -- nominal value of RACT control efficiency

$(\varepsilon_2 - \varepsilon_1)_C$ -- nominal value of $(\varepsilon_2 - \varepsilon_1)$

$(\varepsilon_2 - \varepsilon_1)_T$ -- true value of $(\varepsilon_2 - \varepsilon_1)$

SECTION 9

$\hat{b}_r$ -- relative random uncertainty in emission factor

TLV -- threshold limit value

APPENDIX B

a -- uncertainty associated with A

A -- nominal value of x_1 or $\bar{x}_1$

ABBREVIATIONS AND SYMBOLS (continued)

a_1, a_2 -- lower and upper systematic uncertainties associated with A

b -- uncertainty associated with B

B -- nominal value of x_2 or $\bar{x}_2$

b_1, b_2 -- lower and upper systematic uncertainties associated with B

c -- uncertainty associated with C

C -- nominal value of y or $\bar{y}$

C_1, $C_2 \ldots C_n$ -- constants

d_{n-1} -- Behrens-Fisher statistic

f -- effective degrees of freedom

f'(x) -- first derivative of f(x)

$f(x_1, x_2)$ -- arbitrary function of x_1 and x_2

k, k_α -- constants

n_A, n_B -- sample sizes

$$\mathrm{sgn}(x) \text{ -- } \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \end{cases}$$

s_p -- pooled estimate of standard deviation

s_x -- sample standard deviation of x

$t_{1-\alpha/2,\,\nu}$ -- $(1 - \alpha/2)$ percentage point of t-distribution with ν degrees of freedom

V_A, V_b -- variance factors

x_1, $x_2 \ldots x_n$ -- independent variables

$\bar{x}_1$, $\bar{x}_2$ -- mean values of x_1 and x_2

y -- dependent variable

$\bar{y}$ -- mean value of y

$z_{1-\alpha/2}$ -- $(1 - \alpha/2)$ percentage point of standard normal distribution

α -- one minus confidence level

α_1, α_2, β_1, β_2 -- factors defined in Table A-3

ε_A, $\varepsilon_B \ldots \varepsilon_N$, ε_{AB} -- errors associated with A, B. . .N, and AB

$(\varepsilon_{AB})_{max}$, $(\varepsilon_{AB})_{min}$ -- maximum and minimum values of error in product AB

ABBREVIATIONS AND SYMBOLS (continued)

σ^2_x -- variance of x

APPENDIX C

A -- area of annular region, m^2

A' -- area of sector, m^2

a, b -- parameters in relationship for σ_y

c, d, f -- parameters in relationship for σ_z

D_p -- population density, persons/m^2

F -- "acceptable" concentration, g/m^3

H -- effective emission height, m

P -- affected population, persons

PAAQS -- primary ambient air quality standard, g/m^3

Q -- mass emission rate, g/s

S -- source severity

t -- averaging time, min

TLV -- threshold limit value, g/m^3

t_o -- short-term averaging time, 3 min

u -- wind speed, m/s

x -- downwind distance from emission source, m

x_1, x_2 -- roots of equation $\overline{\chi}/F = 1.0$, m

y -- horizontal distance from plume centerline, m

σ_y -- lateral dispersion coefficient, m

σ_z -- vertical dispersion coefficient, m

χ -- short-term (3 min) time-averaged ground level concentration, g/m^3

χ_{max} -- short-term maximum ground level concentration, g/m^3

$\overline{\chi}_{max}$ -- maximum time-averaged ground level concentration, g/m^3

APPENDIX D

C -- estimated value of effective emission height, m

C_p -- heat capacity at constant pressure of effluent, kcal/g-°K

$\hat{c}_{s\ell}$ -- $c_{s\ell}/(H_s + \Delta H_{calc})$

ABBREVIATIONS AND SYMBOLS (continued)

c_{su}, $c_{s\ell}$ -- upper and lower systematic error bounds for effective emission height, m

$\hat{c}_{su}$ -- $c_{su}/(H_s + \Delta H_{calc})$

D -- stack diameter, m

H_s -- physical stack height, m

M -- molecular weight of effluent, g/g mole

P -- atmospheric pressure

R -- gas constant = 8.314×10^5 dyne-m/g mole-°K

T_a -- ambient air temperature, °K

T_s -- stack gas temperature, °K

u -- wind speed, m/s

V_s -- stack gas exit velocity, m/s

ΔH -- plume rise, m

ΔH_{calc} -- calculated value of plume rise, m

ΔH_{true} -- true value of plume rise, m

APPENDIX E

b - exponent in averaging time correction factor

C -- pollutant exposure concentration, g/m^3

D -- estimated value of "acceptable" concentration, g/m^3

d_ℓ, d_u -- lower and upper uncertainties associated with D, g/m^3

F -- "acceptable" concentration, g/m^3

G -- conversion factor

LD_{50} -- dosage which results in mortality to 50% of exposed population, g/m^3 or $\mu g/m^3$

P -- probability of a lifetime response to pollutant exposure

P* -- "acceptable" probability of a lifetime response to pollutant exposure

PAAQS -- primary ambient air quality standard, g/m^3 or $\mu g/m^3$

$PAAQS_{24}$ -- estimated 24-hr primary ambient air quality standard, g/m^3 or $\mu g/m^3$

t_{avg} -- averging time for PAAQS, hr

TLV -- threshold limit value, g/m^3 or $\mu g/m^3$

ABBREVIATIONS AND SYMBOLS (continued)

PAAQS -- primary ambient air quality standard, g/m^3 or $\mu g/m^3$

$PAAQS_{24}$ -- estimated 24-hr primary ambient air quality standard, g/m^3 or $\mu g/m^3$

t_{avg} -- averging time for PAAQS, hr

TLV -- threshold limit value, g/m^3 or $\mu g/m^3$

APPENDIX F

$\hat{b}_r$ -- relative random uncertainty associated with emission factor

$\hat{e}_r$ -- relative random uncertainty associated with S_C

e_{su}, $e_{s\ell}$ -- upper and lower systematic uncertainties associated with S_C

$\hat{e}_{su}$, $\hat{e}_{s\ell}$ -- upper and lower relative systematic uncertainties associated with S_C

$\hat{e}_u$, $\hat{e}_\ell$ -- upper and lower total relative uncertainties associated with S_C

$F(z)$ -- area under normal curve between minus infinity and z

P -- probability

S_C -- calculated source severity

S_T -- true source severity

z -- standard normal deviate

σ -- standard deviation

μ -- mean

APPENDIX G

B_{meas}, B_{true} -- measured and true values of emission factor, g/kg

$\hat{b}_{su}$, $\hat{b}_{s\ell}$ -- upper and lower relative systematic uncertainties in emission factor

β^+, β^- -- bounds on positive and negative bias in sampling and analysis

APPENDIX H

A -- coefficient in expression for σ_y

A_1, B_1, C_1 -- parameters in expression for σ_z

% Be -- fraction of beryllium in coal

CONS -- coal consumption, g/s

ABBREVIATIONS AND SYMBOLS (continued)

d -- stack diameter, m

EF -- emission factor, g/kg

h -- physical emission height, m

H -- $h + \Delta h$ = effective emission height, m

P -- $\chi_{peak}/\chi_{24\text{-hr}}$

P_b -- barometric pressure, mb

Q -- mass emission rate, g/s

S -- source severity

$\bar{S}_x$ -- mean source severity

$S_{0.95}$ -- 95% point of source severity distribution

T_a -- ambient temperature, °K

TLV -- threshold limit value, g/m^3

T_s -- stack gas temperature, °K

u -- wind speed, m/s

V_s -- stack gas exit velocity, m/s

x -- downwind distance, m

Δh -- plume rise, m

σ_y -- lateral dispersion coefficient, m

σ_z -- vertical dispersion coefficient, m

χ, χ_{peak} -- short-term average ground level concentration, g/m^3

$\chi_{24\text{-hr}}$ -- 24-hr average ground level concentration, g/m^3

APPENDIX I

$\hat{a}_r$ -- relative random uncertainty in average production capacity

$\hat{b}_r$ -- relative random uncertainty in emission factor

$\hat{b}_{s\ell}$, $\hat{b}_{su}$ -- lower and upper relative systematic uncertainties in emission factor

CO -- carbon monoxide

$\hat{c}_r$ -- relative random uncertainty in average emission height

$\hat{c}_{s\ell}$, $\hat{c}_{su}$ -- lower and upper relative systematic uncertainties in average emission height

$\hat{d}_{s\ell}$, $\hat{d}_{su}$ -- lower and upper relative systematic uncertainties in "acceptable" concentration

ABBREVIATIONS AND SYMBOLS (continued)

$(F_C)_i$ -- "calculated" value of "acceptable" concentration for pollutant i, g/m^3
$(F_T)_i$ -- "true" value of "acceptable" concentration for pollutant i, g/m^3
G -- factor which converts TLV to "equivalent" PAAQS
$\overline{G}$ -- average TLV conversion factor
G_i -- TLV conversion factor for pollutant i
m -- 97.5% point of distribution of γ
NO_x -- nitrogen oxides
PAAQS -- primary ambient air quality standard, g/m^3
s -- sample standard deviation
SO_x -- sulfur oxides
t -- averaging time, min
TLV -- threshold limit value, g/m^3
$(TLV)_i$ -- threshold limit value of pollutant i, g/m^3
t_o -- short-term averaging time, min
$\bar{x}$ -- sample mean
β^+, β^- -- bounds on positive and negative bias in sampling and analytical procedures
γ -- correction factor for Gaussian plume equation
μ -- population mean
σ -- population standard deviation
$\overline{\chi}_{max}$ -- maximum time-averaged ground level concentration, g/m^3
$\overline{\chi}_{meas}$, $\overline{\chi}_{pred}$ -- measured and predicted values of ground level concentration, g/m^3

APPENDIX J

$\hat{a}_r$ -- relative random uncertainty in average production capacity
$\hat{b}_r$ -- relative random uncertainty in emission factor
$\hat{b}_{s\ell}$, $\hat{b}_{su}$ -- lower and upper relative systematic uncertainties in emission factor
$\hat{c}_r$ -- relative random uncertainty in emission height
e_ℓ, e_u -- lower and upper uncertainties in source severity
POM -- polycyclic organic material

ABBREVIATIONS AND SYMBOLS (continued)

S_C -- calculated source severity

S_C^* -- critical test value of calculated source severity

S_T -- true source severity

TLV -- threshold limit value

SECTION 1

INTRODUCTION

The Industrial Environmental Research Laboratory (IERL) of EPA uses emissions data to generate information essential to environmental decision making. These data are always subject to a degree of uncertainty because of inherent errors in the methods by which they are obtained. When projections based on the data are employed in decision making, this uncertainty, which extends into the projections, must be considered.

This report develops a decision analysis procedure for dealing with the uncertainty in source assessment information in a systematic and logical fashion. The procedure, as developed, makes allowance, on a mathematically sound basis, for the degree of uncertainty introduced into environmental decision making by uncertainty in the data.

The procedure described in this report applies specifically to the evaluation of technical information used in determining the need to reduce emissions from a source. It does not, however, take into account value judgments, bias in projecting biological tests from test animals to humans, or unpredictable variations in magnitude of emissions due to process upsets, spills, or accidents.

The decision analysis procedure developed in this report constitutes an approximation to standard statistical methods. Such an approximation is necessary, in general, due to the complexity of the types of problems considered. The procedure is simple and straightforward, requiring no detailed knowledge of statistical distribution theory. Application of the decision analysis method requires only the manipulation of a few simple formulas (tabulated in Appendix B) with which most engineers and scientists have some familiarity. Therefore, lack of training in statistical methods should not prove a deterrent to utilization of this procedure.

The decision analysis procedure is illustrated by means of four examples related to environmental control. The first two examples relate to decision problems associated with IERL's Source Assessment Program involving air emissions. The third example involves compliance with a given emissions offset policy concerning industrial plant expansions. The fourth example involves a comparison between two alternative emission control techniques.

The general decision analysis procedure is outlined in Section 3 of the report. The approximations involved in the procedure and its relationship to standard statistical methods are described. In Section 4, the decision indices employed in the Source Assessment Program are defined and their use in decision analysis is discussed.

Sections 5 and 6 present applications of the general decision analysis procedure to two indices used in the Source Assessment Program, namely: source severity and national emissions burden. Source severity combines emissions data, dispersion modeling, and health effects information to yield a measure of the pollution hazard represented by a given source of emissions.

Section 7 addresses a problem involving compliance with local air pollution regulations. The situation considered is that in which an industrial plant expansion must be accomplished without an increase in total pollutant emissions in order to prevent deterioration of ambient air quality. Section 8 considers a problem associated with environmental improvement by means of reasonably available control technology (RACT). The decision to be made is whether to require adoption of RACT in preference to an existing, installed control device.

The major results presented in the report are summarized in Section 2. Conclusions and recommendations are presented in Section 9.

The decision analysis procedure utilizes standard concepts of statistical hypothesis testing. However, there are several areas where the report used a robust application of statistical theory in order to make the decision analysis procedure useful to decision makers who are unfamiliar with detailed statistical methods. Two aspects of the test of hypothesis as used in this report that are not commonly employed in statistical analyses involve: 1) treatment of systematic errors separately from random errors, and 2) estimation of the risks of making incorrect decisions. Those who rigorously apply statistical methods in a puristic fashion may take exception to the approximate statistical methods employed in the report. The discussion presented below is designed to identify areas where approximate statistical methods have been employed.

Treatment of Systematic Errors

Systematic errors in data are treated separately from random errors rather than being combined with them. This treatment is designed to make the systematic errors (bias) more visible to the decision maker so that he can deal with the uncertainty attributed to systematic errors in a conscious fashion. This treatment leads to a family of operating characteristic (OC) curves for the test of hypothesis and to the associated concept of an "effective" OC curve for the test. As explained in Section 5, the "effective" OC curve represents an entire family of OC curves and contains a horizontal section, or plateau, in the center. While such an OC curve may be anathematic to statisticians, this procedure permits systematic errors to be treated in a logical and straightforward manner without recourse to arguments involving statistical distributions. It permits the distinction between bias and imprecision to be maintained throughout the analysis. In addition, the authors believe that this procedure is at least as valid as one which treats systematic and random errors as nonseparable entities.

Estimating the Risk of an Incorrect Decision

In developing the estimated risks of making an incorrect decision, the decision alternatives were handled in a conservative fashion. Normal decision analysis methods divide the decision into three regions: the acceptance region; the indifference region; and the rejection region. Further, usual practice in decision analysis methods is to treat the indifference region as though it is in the acceptance region. The conservative approach taken in the report treats the indifference region as though it is in the rejection region. This approach provides the decision maker opportunity to ensure that a decision is not counter to his charter. Specifically, it enables IERL the opportunity to ensure that pollution control technology is available to meet the requirements of environmental legislation. The decision analysis technique does not and is not intended to ensure that pollution control technology is applied. The decision to apply control technology requires input for trade-off (economic, social, political, etc) considerations to be made and those considerations are beyond the scope of IERL's source assessment program.

SECTION 2

SUMMARY

BACKGROUND

The objective of completing a source assessment in the Source Assessment Program is to provide sufficient information to aid IERL in determining the need for reducing pollution from stationary sources. A source is an entire industrial, commercial, or municipal operation which is national in scope. An assessment is the evaluation of a source's pollution based on all available information on process emissions, discharges, and pollution control. The product of a source assessment is a Source Assessment Document. The ultimate result is an EPA decision regarding the need for further study of the source.

This report describes a decision analysis procedure that can be used as an aid in evaluating the information contained in Source Assessment Documents. This procedure, which constitutes an approximate statistical analysis, employs the concepts of error propagation, confidence intervals, and hypothesis testing. The statistical concepts cannot be applied in a rigorous manner due to both the complexity of the problems considered and the subjective nature of environmental decision making.

The general decision analysis principles are illustrated by four examples in the areas of emissions assessment and control describdescribed in the following subsections.

SOURCE SEVERITY

Source severity is defined as the ratio of two concentrations: 1) concentration of pollutant to which the human population may be exposed and 2) concentration of pollutant which represents an "acceptable" concentration. The exposure concentration is the time-averaged maximum ground level pollutant concentration as determined by Gaussian plume dispersion methodology. The "acceptable" concentration is estimated in two ways. It is the primary ambient air quality standard for criteria pollutants, or it is a surrogate primary ambient air quality standard for noncriteria pollutants.

The values of source severity which IERL uses as a guide for screening sources are shown in Table 1. The value of $S = 0.05$ as

TABLE 1. MEASURE/ANALYSIS ASPECTS OF SOURCE SEVERITY

Measure[a,b,c]	Analysis
$S \equiv \frac{C}{F} \geq 0.05$	There is sufficient cause to consider a source as a candidate for further study.
$S \equiv \frac{C}{F} < 0.05$	There is not sufficient cause to consider a source as a candidate for further study.

[a] S = calculated source severity.

[b] C = concentration to which the population is exposed.

[c] F = potentially hazardous concentration.

a milepost was obtained by evaluating the uncertainties in sampling and analysis results, atmospheric dispersion models, and health effects data.

NATIONAL EMISSIONS BURDEN

The national emissions burden is the mass of each criteria pollutant emission from a source divided by the national mass of each criteria pollutant emission. This index uses engineering and emissions data for a source to develop an emissions inventory for that source. The values which IERL uses as a guide for screening sources are shown in Table 2. The value of $B_N = 0.05$ as a milepost was obtained by evaluating the uncertainties in sampling and analysis data and the uncertainties in national emission inventories such as those generated by EPA's National Emissions Data System (NEDS). In short, if emissions from a source amount to more than 0.05% of the U.S. total emission rate for a given criteria pollutant, the source is considered as a candidate for further study.

OFFSET CALCULATIONS

As part of EPA's overall program to achieve and maintain national ambient air quality standards, the states are required (through state implementation plans) to determine the amount of emissions reduction necessary to offset the probable impact of increased population, industrial activity, motor vehicle traffic, and other growth factors. Of particular concern is the policy regarding increased emissions due to increased industrial activity through expansions of existing plants.

One way to offset emissions from a plant expansion is to require the plant to make a corresponding emissions reduction from its existing unit. Such a requirement entails the associated problem of determining *a posteriori* whether the plant is complying with

TABLE 2. MEASURE/ANALYSIS ASPECTS OF THE NATIONAL EMISSIONS BURDEN

Measure[a,b,c]	Analysis
$B_N \equiv \frac{\Sigma M_P}{\Sigma M_N}(100) \geq 0.05$	There is sufficient cause to consider a source as a candidate for further study.
$B_N \equiv \frac{\Sigma M_P}{\Sigma M_N}(100) < 0.05$	There is not sufficient cause to consider a source as a candidate for further study.

[a] B_N = national criteria pollutant emissions burden, percent.

[b] M_P = mass of criteria pollutant emissions from a source.

[c] M_N = national mass of criteria pollutant emissions.

the offset policy. Such a determination is not entirely straightforward because the emission rates are not known exactly. This uncertainty should be taken into account in the decision analysis process.

To quantify the uncertainty in the emission rates before and after plant expansion, error bounds on the following quantities are required:

- Emission rates from each emission point in the new unit.
- Emission rates before and after expansion from each emission point in the original unit which has different emissions after expansion.

A simple, approximate statistical test is used to determine whether the plant may be in compliance with the offset policy. In order to ensure that air quality is not degraded, the procedure requires that emissions from the original plant be reduced by an amount sufficient to compensate not only for the emissions from added capacity, but also for the uncertainty in emissions data.

REASONABLY AVAILABLE CONTROL TECHNOLOGY

In order to meet environmental regulations, a plant may be required to install reasonable available control technology (RACT) based on a comparison of two alternative control technologies. For this analysis, it is assumed that a plant currently employs a given control technology having a nominal (measured) control efficiency which is less than that of RACT for a given application. In order that the plant be required to adopt RACT, it should be established that RACT is indeed superior to the installed control technology. Since neither of the two control efficiencies can be known with certainty, this determination should not be based on the nominal efficiencies alone.

A simple, approximate statistical test can be employed to compare the effectiveness of RACT with that of the installed technology. The essence of the procedure is to ensure that the incremental emissions reduction obtained by RACT more than compensates for the uncertainties in the control efficiencies for the existing control device and RACT.

SECTION 3

STATISTICAL BASIS FOR THE PROCEDURE

The general procedure set forth in this report is based on the statistical concepts of confidence intervals and hypothesis testing. In this initial effort, no attempt has been made to apply alternative decision-theoretic precepts such as the minimax principle or the principle of maximum expected utility (1, 2).

The general procedure can be summarized as follows:

- Define a numerical index to be used as an aid in decision making. Source severity, defined in Section 4, is such an index. The value of the index may depend upon experimental data, a mathematical model, and/or a semiempirical relationship, each of which introduces an element of uncertainty into the problem.
- Perform an error analysis to determine the uncertainty in the value of the index due to uncertainties in the data, the mathematical model, etc.
- Based on the results of the error analysis, formulate a statistical test of hypothesis as the basis for using the index in the decision-making process.
- Set the values of controllable parameters (e.g., the level of the test, the critical test value, the acceptable random uncertainty in the data) to keep the probability of making an incorrect evaluation of the index within acceptable limits.

The above procedure may be an iterative one. For example, after completing the above steps, the decision index may be redefined in order to reduce the risk of incorrect decisions. In this manner, the entire decisionmaking process, from initial program planning through data collection to the final decision, is focused toward and guided by the needs of the decision maker. Specifically, the decisionmaking procedure is based on the concept of controlling the risk of making an incorrect decision.

---

(1) Savage, L. J. The Foundations of Statistics. John Wiley & Sons, Inc., New York, New York, 1954.

(2) Lindley, D. V. Making Decisions. John Wiley & Sons, Ltd., London, 1971.

At the heart of the above procedure is the second step--error analysis. This analysis makes use of formulas for the propagation of random and systematic errors presented in Appendix B. The error analysis procedure can be summarized as follows:

- Use formulas for propagation of random errors to obtain the random component of uncertainty in the index.
- Use formulas for the propagation of systematic errors to obtain the systematic component of uncertainty in the index.
- Combine the results of the above two steps to obtain the total uncertainty in the index; i.e., the interval of uncertainty associated with the index.

The above error analysis procedure represents an approximate statistical technique, as shown in Appendix B. The last two steps of the procedure, involving the test of hypothesis, employ standard statistical procedures. However, since the test of hypothesis is based on the results of the error analysis, the entire procedure constitutes an approximate statistical technique.

The nature of the above approximation and the circumstances which necessitate it are discussed in the following paragraphs. The discussion involves the calculation of confidence intervals using the "Student t" statistic. This subject is covered in most books on elementary statistics, such as Reference 3.

Let I_C denote the ("calculated") value of the decision index obtained from experimental data, and let I_T denote the ("true") value which would be obtained if there were no errors in the data, the mathematical model employed, etc. I_C can be regarded as a value assumed by a random variable, $\underline{I_C}$, which is an estimator or the parameter I_T.[a] The value, I_C, is referred to as a point estimate of I_T, and the simplest approach to decision making is to base the decision on this value. However, in order to take account of the uncertainty in the data and thereby control the risk of making an incorrect decision, it is desirable to obtain an interval estimate (i.e., a confidence interval) for I_T.

The standard statistical approach to this problem is to determine the probability distribution of the random variable $\underline{I_C}$, and to use this distribution to construct a confidence interval for I_T and/or to set up a statistical test of hypothesis concerning I_T. The confidence interval thus obtained or, alternatively, the outcome of the statistical test, then determines the decision to be made.

---

[a] Underbars denote random variables.

----

(3) Freund, J. E. Mathematical Statistics. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962. 390 pp.

The above statistical approach is not generally applicable to the type of problem considered in this report for the following reasons:

- The estimator, $\underline{I_C}$, is in general biased due to biases in the experimental methods, and the value of the bias in $\underline{I_C}$ is not known.
- $\underline{I_C}$ is generally a complex function of other random variables, so that its distribution function cannot be determined analytically.

The approach used in this report utilizes the principles of error propagation to obtain an approximate confidence interval for I_T. Although an approximation, the resulting confidence interval is generally conservative, thereby permitting the establishment of an upper bound on the risk of an incorrect decision.

To aid in clarifying the relationship between the approximate procedure used in this report and the rigorous statistical approach, two simple examples are given below. Both examples involve variations of the fundamental problem of estimating the mean, μ, of a normally distributed population having variance, σ^2, given a random sample of size n from that population. In the usual procedure, the sample mean, $\bar{X}$, and sample standard deviation, s, are used to obtain a confidence interval for μ:

$$\mu = \bar{X} \pm \frac{ts}{\sqrt{n}} \tag{1}$$

where t represents the appropriate value of Student's statistic.

Example A

Suppose now that instead of $\bar{X}$, we use a different statistic, $\hat{X}$, to estimate the mean, μ. Further, suppose the random variable $\underline{\hat{X}}$ has the same distribution as $\underline{\bar{X}}$, except that $\underline{\hat{X}}$ is biased by an amount, b. That is, the expected value of $\underline{\hat{X}}$ is given by:

$$E(\underline{\hat{X}}) = \mu - b \tag{2}$$

We would then obtain a confidence interval for μ as follows:

$$\mu - b = \hat{X} \pm \frac{ts}{\sqrt{n}} \tag{3}$$

$$\mu = \hat{X} + b \pm \frac{ts}{\sqrt{n}} \tag{4}$$

$$\hat{X} + b - \frac{ts}{\sqrt{n}} \le \mu \le \hat{X} + b + \frac{ts}{\sqrt{n}} \tag{5}$$

Suppose, however, that we do not know the value of b, but that we have bounds on it as follows:

$$- b_{\ell} \leq b \leq b_u \tag{6}$$

Combining Inequalities 5 and 6 gives:

$$\hat{X} - b_{\ell} - \frac{ts}{\sqrt{n}} \leq \mu \leq \hat{X} + b_u + \frac{ts}{\sqrt{n}} \tag{7}$$

Inequality 7 does not represent a confidence interval in the usual sense, because we cannot associate a specific confidence level with it. However, if Inequality 6 is strictly true, then Interval 7 contains Interval 5. Hence, the confidence level associated with Interval 7 must be at least as great as that associated with Interval 5. For example, if the latter represents a 95% confidence interval, the confidence level associated with the former must be greater than or equal to 95%. Furthermore, Inequality 7 is the smallest interval that could be constructed with the given information for which the confidence level is at least 95%.

In the approximate procedure used in this report, formulas for the propagation of systematic errors are used to obtain bounds on the bias as in Inequality 6. These bounds are then used to construct an approximate confidence interval as in Inequality 7.

Example B

Consider next another variation of the standard problem in which it is desired to estimate the sum of means from several normal populations based on independent random samples from each of these populations. (For simplicity, we consider the sum of two means only.) Thus, we wish to estimate $(\mu_1 + \mu_2)$ based on samples of size n_1 and n_2 from two normal populations whose variances, σ_1^2 and σ_1^2, are unknown. In this case, there is no generally applicable statistic corresponding to the t statistic which can be used to obtain a confidence interval for $\mu_1 + \mu_2$. The method reported here uses formulas for the propagation of random errors to generate an approximate confidence interval which is based on a weighted average of the two individual t-values. The (approximate) interval corresponding to Interval 1 is:

$$(\mu_1 + \mu_2) = (\overline{X}_1 + \overline{X}_2) \pm t_{ave}\, s \tag{8}$$

where

$$s = \left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^{1/2} \tag{9}$$

$$t_{ave} = \left(f_1 t_1^2 + f_2 t_2^2\right)^{1/2} \tag{10}$$

$$f_1 = \frac{s_1{}^2/n_1}{s_1{}^2/n_1 + s_2{}^2/n_2} \tag{11}$$

$$f_2 = \frac{s_2{}^2/n_2}{s_1{}^2/n_1 + s_2{}^2/n_2} \tag{12}$$

In using the error propagation formulas, Interval 8 is obtained directly without the necessity of computing either t_{ave} or s.

In the general case, the problems illustrated in examples A and B both occur, and it is necessary to use both random and systematic error propagation formulas. The two calculations are made separately, and the results are then combined to obtain an approximate confidence interval. This interval can be used directly or, alternatively, it can be used to set up an approximate statistical test for decision making.

SECTION 4

SOURCE ASSESSMENT METHODOLOGY

DEFINITION OF INDICES

The study of air emissions in the Source Assessment Program comprises a screening procedure designed to identify the need for application of or development of new control technology for stationary sources of air pollution. Five numerical indices are used in the program to identify those industries, or source types, which are considered to be potential candidates for emissions reduction.

Source Severity

Source severity, S_C, is defined as

$$S_C \equiv \frac{\overline{\chi}_{max}}{F} \tag{13}$$

where $\overline{\chi}_{max}$ is the maximum time-averaged ground level concentration of a given pollutant emitted from the source type in question, and F represents an "acceptable" concentration of that pollutant. The source type is considered to be a potential candidate for emissions reduction if $S \geq 1.0$. Since the true value of severity is unknown, however, the decision must be based on an estimated, or "calculated," value of source severity, S_C. In recognition of the uncertainty associated with the estimated severity, the cut points for decision making were originally (June, 1975) set up as follows:

If $S_C \geq 1.0$:	Source type is a potential candidate for emissions reduction.
If $0.1 \leq S_C < 1.0$:	Source type may be a potential candidate for emissions reduction.
If $S_C < 0.1$:	Source type is not a potential candidate for emissions reduction.

National Emissions Burden

The national emissions burden is defined as the ratio M_p/M_n where M_p denotes the total annual mass of emissions of each criteria pollutant from a given source type, and M_n denotes the total annual mass of emissions of that pollutant from all stationary sources nationwide. When the national emissions burden from a specific source type exceeds 0.001 (0.1%), the source type is considered to be a potential candidate for emissions reduction.

State Emissions Burden

The state emissions burden is defined as the ratio M_p/M_s where M_p now denotes the annual mass of emissions of each criteria pollutant from a given source type within a given state, and M_s is the annual mass of emissions of that pollutant from all stationary sources within that state. When the state emissions burden exceeds 0.01 (1%), the source type is considered to be a potential candidate for emissions reduction.

The preceding are the three dominant screening criteria and comprise the basic components of the decisionmaking process. Since the first index is directly proportional to the emission rate from a given source, a discrepancy arises when comparing a source type having a few large plants with a source type having many small plants. The second and third criteria, which deal with the total mass of emissions, are designed to overcome this problem.

The two remaining decision indices are used to support or modify decisions based on the first three indices.

Affected Population

Affected population designates the average number of persons exposed to high concentrations (e.g., those for which $\overline{\chi}/F>1.0$) of a given emission from a given source. This quantity is useful in characterizing emissions because a given source may exceed the first three criteria; yet, if it is located in a sparsely populated area, it may have a relatively small impact on human health compared with a source located in a densely populated area. In addition, a source may have a large value of source severity due to a low emission height. Again, its impact on human health may be small because the low emission height results in pollutants being dispersed over a very small area in the immediate vicinity of the source. The calculation of affected population is described in Appendix C.

Growth Factor

The growth factor is determined from the ratio of known to projected emissions from a source type. For example,

$$\text{Growth factor} = \frac{\text{Projected emissions in 1978}}{\text{Emissions in 1973}} \qquad (14)$$

Other 5-yr periods (e.g., 1975 and 1980) could also be used depending on available data. The main purpose of this index is to eliminate from consideration those sources whose emissions are expected to decrease greatly in the near future; e.g., due to the implementation of new emission controls or to a process being phased out of use.

In the subsequent analysis of uncertainty, the discussion centers mainly on the first decision index, source severity. The national emissions burden is also considered in more detail in Section 6.

EXPANDED DISCUSSION OF SOURCE SEVERITY

Basic Definition

As noted previously in Equation 13, the source severity, S_C, for a given pollutant emitted from a given source is defined as

$$S_C \equiv \frac{\bar{\chi}_{max}}{F}$$

where $\bar{\chi}_{max}$ equals the maximum time-averaged ground level concentration of pollutant which would result from the source emitting into a standard receiving atmosphere, and where F equals an "acceptable" concentration. The standard receiving atmosphere is defined as one having Class C stability and a mean wind speed of 4.5 m/s. The averaging time for the mean concentration is the same as that used in the specification of F. For the Source Assessment Program, the "acceptable" concentration is specified as follows:

F = PAAQS[a] for criteria pollutants

= (TLV)G for noncriteria pollutants (4)

where G is a factor which converts the TLV into an "equivalent" PAAQS. The averaging time used in the PAAQS for each criteria pollutant is shown in Table C-4 of Appendix C. The averaging time used for noncriteria pollutants is 24 hr.

Appendix E develops a correlation between the Primary Ambient Air Quality Standard for criteria pollutants and their corresponding TLV's. The analysis indicates that "G" for criteria pollutant TLV's ranges from 0.0260 to 0.158 with an average value of 0.0467.

---

[a] PAAQS represents primary ambient air quality standard.

- - - - - - - - - - -

(4) TLVs® Threshold Limit Values for Chemical Substances and Physical Agents in the Workroom Environment with Intended Changes for 1976. American Conference of Governmental Industrial Hygienists, Cincinnati, Ohio, 1976. 94 pp.

Calculation of Source Severity

Ideally, S_C would be determined by placing the source in the standard receiving atmosphere and measuring the resulting maximum ground level concentration. If the measurement could be made with no error, the measured value would yield the true source severity. Obviously, such a procedure is impossible in practice. Therefore, S_C must be estimated by measuring the source emission rate and by using dispersion modeling to relate emission rate to $\bar{\chi}_{max}$. In the Source Assessment Program, the simple Gaussian plume equation is used to calculate source severity as follows (5, 6):

$$S_C = \frac{2\ Q}{\pi e u F H^2}\left(\frac{3}{t}\right)^{0.17} = \frac{2\ CAP\ EF}{\pi e u F H^2}\left(\frac{3}{t}\right)^{0.17} \qquad (15)$$

where Q = mass emission rate, g/s
H = effective emission height, m
u = average wind speed (4.5 m/s)
e = natural base logarithm $\cong$ 2.72
CAP = production capacity, kg/s
EF = emission factor; i.e., mass of emissions generated per unit of product produced, g/kg
t = averaging time for mean concentration, min

The concentration obtained from the Gaussian plume equation corresponds to an averaging time of approximately 3 min (5). The factor $(3/t)^{0.17}$ in Equation 15 corrects this value for averaging times between 3 min and 24 hr (5, 7). Equation 15 is used for all pollutants with the exception of nitrogen oxide (NO_x), for which the primary standard averaging time is 1 yr. Since the above correction factor is not valid for averaging times of this magnitude, a modified approach, which is described in Appendix C is necessitated.

---

(5) Turner, D. B. Workbook of Atmospheric Dispersion Estimates. Public Health Service Publication No. 999-AP-26, U.S. Department of Health, Education, and Welfare, Cincinnati, Ohio, 1969. 84 pp.

(6) Pasquill, F. Atmospheric Diffusion, Second Edition. John Wiley & Sons, Inc. (Halsted Press), New York, New York, 1974. 429 pp.

(7) Cheremisinoff, P. N., and A. C. Morresi. Predicting Transport and Dispersion of Air Pollutants from Stacks. Pollution Engineering, 9(3):3, 26, 1977.

The conversion factor, G, for converting TLV to an equivalent PAAQS is specified as follows:

$$G = \frac{1}{300} = \frac{8}{24}\,\frac{1}{100} = 0.0033$$

Theoretically, the factor 8/24 adjusts the TLV from an 8-hr work day to continuous (24-hr) exposure, and the factor of 1/100 is designed to account for the fact that the general population constitutes a higher risk group than healthy workers. Alternative methods of estimating "acceptable" concentrations for non-criteria pollutants are considered in Appendix E.

REPRESENTATIVE SOURCE

In assessing the potential pollution problems associated with an industry or source type, it is necessary to take into account plant-to-plant variations in physical parameters such as stack height, production capacity, or process technology. In the Source Assessment Program, this problem is handled in two ways: 1) through the concept of a representative source and 2) through the use of source severity distributions. The representative source concept is discussed in this subsection; source severity distributions are discussed in the following subsection.

The concept of a representative source constitutes a simplified approach to plant-to-plant variability in which an industry or source type is characterized by means of a "typical" (or representative) source or sources. In general, the representative source is a hypothetical plant having physical parameters which are typical or average values for the source type in question. The methodology employed in defining a representative source is described in the following discussion (8).

Simple Case

The representative source, for a simple case, can be defined as one having an average plant capacity, average stack height, average county population density, and an average emission factor.

The following definitions can be used in calculating required average parameters.

- Average plant capacity is the total capacity of the industry divided by the number of plants.
- Average stack height is the sum of all known stack heights for that industry divided by that number of stacks.

---

(8) Hughes, T. W., R. B. Reznik, R. W. Serth, and Z. S. Kahn. Source Assessment Methodology. Contract 68-02-1874, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, January 7, 1976. 33 pp.

- Average county population density is the sum of the population densities of all counties within which plants are located divided by the number of sites.

- Average emission factor is the sum of the known emission factors divided by the number of factors known.

These definitions are valid when there is no correlation between plant size, emission factor, stack height, and population density and when emission factor data are taken from a random sample of the population. One simple way to detect possible correlations is to graphically depict the variables in question; e.g., plant size versus population density.

Special Cases

Small Number (Fewer than 10) of Sources--

When a small number (fewer than 10) of sources comprises the source type being studies, a representative source is defined as the largest source within the group. It may also be desirable to treat each particular source individually.

Large Number of Sources--

For source types with a large number of sites, it is impractical to determine all of the average parameters. This problem can be handled in several ways depending on the data base available. Some possibilities include

- Consideration of 30 plants at random. If the parameters are normally distributed, this is sufficient to find an average. Random election of 30 plants will give fairly accurate estimates of the industry's mean parameters while describing the variation within the industry.

- Use of a simulation technique.

- Use of state average population density.

Nonhomogeneous Sources--

When qualitative differences in subgroups exist, a representative source is defined by subdividing into homogeneous subgroups and considering each group separately.

Log-Normal Distributions--

In some sources, the plant capacity is log-normally distributed. Most of the plants are small, but the few large plants account for a major portion of the production. Under these circumstances, it would be better to subdivide into size groups or to treat only the larger plants as a worst case condition.

Correlation of Averaging Parameters--
Sources which exhibit a correlation between the averaging parameters (emission factor, plant size, stack height, population density) should be treated on a case-by-case basis. If the emission factor or population density depends on plant size, a weighted average should be calculated. For other situations, calculating the distribution of severities using a simulation technique may be appropriate.

SOURCE SEVERITY DISTRIBUTIONS

A source severity distribution depicts the distribution of source severities among individual plants within an industry. It thus provides more detailed information about potential air pollution problems associated with a given industry than does the representative source severity. When sufficient information is available to compute a source severity for each plant in an industry, a deterministic source severity distribution can be constructed for the industry. An example of a deterministic severity distribution is shown in Figure 1 (9). When such detailed information is not available, a simulation technique can be employed to generate an approximate (simulated) source severity distribution based on data from a sample of plants. The methodology used to simulate source severity distributions is described in Reference 10.

STOCHASTIC APPROACH TO SOURCE SEVERITY

In the Source Assessment Program, a deterministic approach to calculating source severity has been adopted by employing the concept of a fixed receiving atmosphere. The main advantage of this approach is its simplicity: It permits a wide variety of source types to be analyzed and compared on a consistent basis with a minimum of experimental and computational effort.

An alternative approach is to allow meteorological parameters (i.e., stability class or wind speed) to vary stochastically. This approach generates a frequency distribution of source severities rather than a single value, as is obtained in the

---

(9) Serth, R. W., and T. W. Hughes. Source Assessment: Carbon Black Manufacture. Contract 68-02-1874, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina. (Preliminary document submitted to the EPA by Monsanto Research Corporation, December 1975.) 145 pp.

(10) Eimutis, E. C., B. J. Holmes, and L. B. Mote. Source Assessment: Severity of Stationary Air Pollution Souces--A Simulation Approach. EPA-600/2-76-032e, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, July 1976. 119 pp.

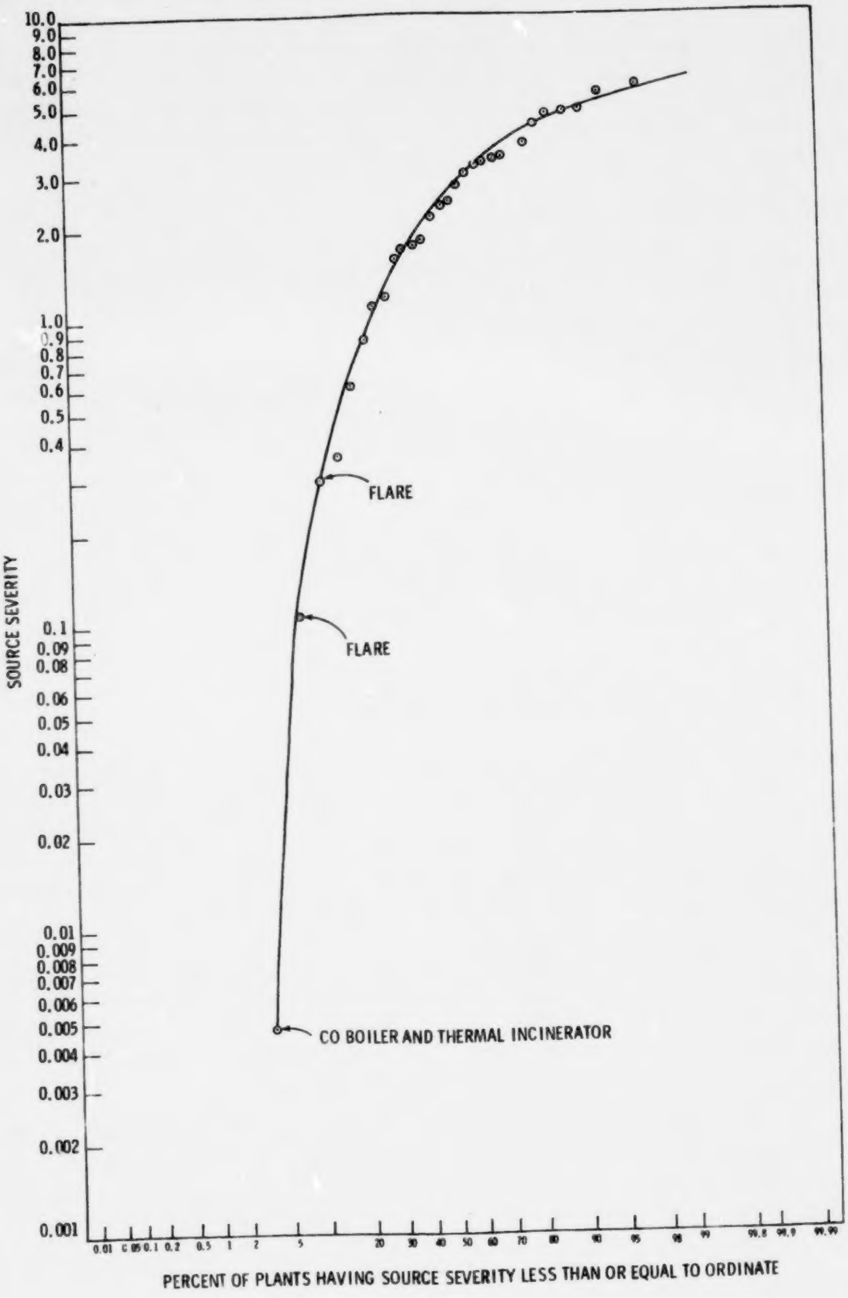


Figure 1. Deterministic source severity distribution for carbon monoxide emissions from carbon black plants (9).

deterministic approach. For making decisions concerning the need for emissions reduction, a value of source severity corresponding to a relatively high (e.g., 99 or 99.9) percentage point of this distribution would be appropriate. For example, for severities based on 24-hr mean concentrations, values higher than the 99% point would be expected to occur about four times per year. Thus, a low value of severity at the 99% point would indicate a low risk of potentially hazardous conditions arising from the source in question. The same would not be true of a low value of the mean severity, for example, since severities much higher than the mean might occur many times over the course of a year.

The stochastic approach to the calculation of source severity is investigated in Appendix H. A Monte Carlo simulation and probabilistic sensitivity analysis are carried out in which wind speed, stability class, lateral and vertical dispersion coefficients, peak-to-mean concentration ratio, etc., are allowed to vary stochastically. The output from the simulation is a probability distribution of values of source severity for beryllium emissions from a coal-fired powerplant. In this example, the upper 95% point of the severity distribution is approximately 1.9, and the upper 99% point is approximately 3.0.

By comparison, the deterministic value of severity (corrected for plume rise) for the above example is 3.3. This value corresponds roughly to the 99% point of the stochastic severity distribution. Thus, on the basis of this example, it appears that the deterministic approach is consistent with the stochastic approach.

The simulation described in Appendix H is based on the generalized Gaussian dispersion equation (Equation H-7). Hence, the results are restricted to meteorological conditions under which the Gaussian model is valid. A complete solution to the problem would have to include dispersion models applicable to other meteorological conditions, such as fumigation (inversion breakup) and trapping, and the frequency with which the various meteorological conditions occur.

Although ground level concentrations under trapping and fumigation conditions can be considerably higher than those predicted by the Gaussian model (coning conditions), these conditions generally persist for relatively short periods (30 min to 4 hr). Therefore, these conditions may have a relatively small effect on 24-hr mean concentrations, which are of primary interest in relation to source severity. The above reasoning provides justification for the exclusive use of the Gaussian model as an approximation in the simulation procedure.

The example given in Appendix H also indicates that the deterministic approach to source severity yields a worst case value in the sense that higher values would be expected to occur only infrequently, on the order of once per year.

SECTION 5

SOURCE ASSESSMENT: SOURCE SEVERITY

INTRODUCTION

In this section, the general procedure outlined in Section 3 is applied to the source severity decision index. It was noted in the previous section that the true value of source severity is always unknown and that decisions must be based on an estimated, or "calculated," value of severity. The uncertainty associated with the estimated value of severity is related to uncertainties in individual parameters (i.e., production capacity or emission factor) used to calculate severity. In the following subsections, a mathematical expression for this relationship is obtained and used to establish overall guidelines for decision making in the Source Assessment Program.

Only the source severity of a representative source will be considered. In addition, it will be assumed that the source type in question is homogeneous with respect to emission factor, so that the representative source emission factor can be obtained by sampling at a single plant.[a] This assumption may appear to pose a severe restriction on the applicability of the analysis; in practice, however, economic constraints often dictate that sampling be conducted at a single plant. Use of the measured emission factor to calculate the severity of the representative source then involves the implicit assumption of homogeneity with respect to emission factor.

For the purpose of this analysis, it is assumed that the true source severity, S_T, is given by the following equation:

$$S_T = \gamma \left(\frac{2\,(CAP)(EF)}{\pi e u F H^2}\right)\left(\frac{t_o}{t}\right)^{0.17} \tag{16}$$

where γ is a correction factor for the simple Gaussian plume equation. Since we are dealing with the representative source, CAP and H now represent arithmetic mean production capacity and effective emission height, respectively, for the source type in question. The emission factor, EF, in Equation 16 is the emission factor for the plant at which sampling was performed.

---

[a] This assumption is made to simplify the analysis and does not represent an inherent limitation of the general procedure.

The following discussion will analyze Equation 16 by defining the uncertainties contained in each of the variables, by showing how the uncertainties are propagated, by developing alternate decision tests (tests of hypothesis), by quantifying and comparing the alternate decision tests, and by developing guidelines for the Source Assessment program. Two alternate decision tests will be developed analytically - a deterministic approach and a stochastic approach. The deterministic approach assumes that the true value of source severity is a fixed parameter and treats large uncertainties in Equation 16 as systematic errors. The stochastic approach assumes that the true value of source severity is a random parameter (i.e., changes daily) and treats all uncertainties as random errors.

UNCERTAINTIES IN PARAMETERS

The uncertainties associated with each of the individual parameters appearing in Equation 16 are discussed below. The concepts of random and systematic uncertainties are discussed in Appendices A and B.

CAP (Average Production Capacity)

If CAP is determined from a sample of plants in the industry or source type, there is a random sampling uncertainty which shall be equated with the 95% confidence interval about the sample mean plant capacity. If CAP is determined from a knowledge of all plants in the industry, then this uncertainty is zero.

H (Average Effective Emission Height)

The average effective emission height, H, is comprised of two components, $H = h+\Delta H$, where h is the physical stack height and ΔH is the plume rise. A statement similar to the one made for CAP applies to h as well. In addition, there is uncertainty in estimating plume rise, ΔH. The error in plume rise can be assumed to be a randomly distributed uncertainty over the population of all sources and atmospheric conditions, provided ΔH is indeed included in H.

In the development of the deterministic decision test later, source severity is computed for a fixed source and a fixed receiving atmosphere.[a] Thus, the estimated severity, S_C, can be considered as a sample value from a population of severities in which both source type and receiving atmosphere are fixed. In this use, the error in plume rise is treated as a bias in S_C since the error is amplified by squaring H in Equation 16. In the development of the stochastic decision test developed later, source severity is computed for a fixed source and a variable receiving atmosphere. In this use, the error in plume rise represents a random uncertainty in S_C.

---

[a] The receiving atmosphere was defined by specifying stability class and wind speed. These two parameters do not suffice to define a unique state of the atmosphere. However, it may be assumed that the additional meteorological parameters necessary to define a unique atmospheric state are maintained at fixed, albeit unspecified, values.

EF (Emission Factor)

There is an uncertainty in EF which is composed of process variation, sampling uncertainty, analysis uncertainty, and uncertainty in production rate during sampling. The total uncertainty has both a random and a systematic component, the latter resulting from possible bias in the sampling and analytical procedures.

The random component is given by:[a]

$$b_r^2 = b_{r1}^2 + b_{r2}^2 + b_{r3}^2 + b_{r4}^2 \tag{17}$$

where b_r = random uncertainty in emission factor, g/kg

b_{r1} = uncertainty in emission factor due to imprecision in sampling, g/kg

b_{r2} = uncertainty in emission factor due to imprecision in analysis, g/kg

b_{r3} = uncertainty in emission factor due to process variation, g/kg

b_{r4} = uncertainty in emission factor due to imprecision in measurement of production rate, g/kg

γ (Correction Factor)

There is an uncertainty in the correction factor, γ, which is a reflection of the uncertainty inherent in the Gaussian dispersion equation itself. For the purpose of this analysis, values of γ are assumed to be log-normally distributed over source types and atmospheric conditions with mean of 1.0 and a variance of 0.264 (standard deviation = 0.514). The 2.5% and 97.5% points of the distribution are taken to be 1/m and m (where m = 3).

In the deterministic decision test developed later, the uncertainty in γ is treated as a systematic uncertainty with m = 3. In the stochastic decision test, the uncertainty in γ is treated as a randam uncertainty with a variance of 0.264.

---

[a] If the source type is not assumed to be homogeneous with respect to emission factor, then the random uncertainty contains an additional term due to plant-to-plant variability of emission factor. In general, the plant-to-plant distribution of emission factors is not a normal distribution, which further complicates the analysis.

$(t_o/t)^{0.17}$ (Averaging Time Correction Factor)

For the purpose of the present analysis, the uncertainty associated with the averaging time correction factor $(t_o/t)^{0.17}$ is assumed to be included in the uncertainty associated with γ. This assumption is discussed further in the following section.

F (Acceptable Pollutant Concentration)

There may be an uncertainty associated with the "acceptable" pollutant concentration, F, depending upon interpretation. It can be argued that there is no pollutant concentration, however small, which is entirely innocuous. Therefore, specification of an acceptable concentration represents a subjective judgment on the part of the decision maker. Hence, the "true" value of F may be regarded as whatever value the decision maker specifies. In this case then, there is no uncertainty in F since the "true" value equals the "calculated" (specified) value.

Alternatively, it can be argued that the PAAQS represents a pollutant concentration at which no adverse effects on human health can be detected. Therefore, the "true" value of F can be equated with the PAAQS. In this case, there is no uncertainty in F for criteria pollutants. For noncriteria pollutants, however, there is uncertainty associated with the conversion factor, G, which converts the TLV into an equivalent PAAQS. This uncertainty can be assumed to be randomly distributed over the population of all pollutants. However, since source severity is computed for a given pollutant, this uncertainty should be treated in the analysis as a systematic uncertainty.

DETERMINISTIC DECISION APPROACH

In developing a deterministic decision test it is assumed that the true value of source severity is a fixed parameter. The test is developed for a fixed source in a fixed atmosphere. It treats small uncertainties as random errors and large uncertainties as systematic errors. The approach employs the development of formulae for the propagation of errors.

Propagation of Random Uncertainties

The total uncertainty in source severity is determined by the propagation of uncertainties in CAP, H, EF, γ, and F through Equation 16. For convenience, let A, B, C, and D represent estimated or measured (as opposed to true) values as follows:

A = CAP
B = EF
C = H
D = F

The "measured" value of γ is taken as unity. Furthermore, let

$$k = \frac{2}{\pi eu}\left(\frac{t_o}{t}\right)^{0.17} = \text{constant} \tag{18}$$

Considering the random uncertainty first,

$$S_C \pm e_r = \frac{k}{D}\frac{\left(A \pm a_r\right)\left(B \pm b_r\right)}{\left(C \pm c_r\right)^2} \tag{19}$$

where a_r, b_r, c_r, and e_r are the random uncertainties in A, B, C, and S_C, respectively, and S_C is the nominal, or "calculated," source severity; i.e.,

$$S_C = \frac{kAB}{DC^2} \tag{20}$$

Thus, if we are working at the 95% confidence level, $(A \pm a_r)$ represents a 95% confidence interval for the industry mean production capacity, CAP.

Now let

$$a_r = \hat{a}_r A$$

$$b_r = \hat{b}_r B$$

$$c_r = \hat{c}_r C$$

$$e_r = \hat{e}_r S_C$$

i.e., $\hat{a}_r$, $\hat{b}_r$, $\hat{c}_r$, and $\hat{e}_r$ are the relative random uncertainties associated with A, B, C, and S_C. Using the error propagation formula for multiplication (see Appendix B),

$$\begin{aligned}\left(A \pm a_r\right)\left(B \pm b_r\right) &= AB \pm \sqrt{B^2 a_r{}^2 + A^2 b_r{}^2} \\ &= AB \pm \sqrt{B^2 A^2 \hat{a}_r{}^2 + A^2 B^2 \hat{b}_r{}^2} \\ &= AB \pm AB\sqrt{\hat{a}_r{}^2 + \hat{b}_r{}^2}\end{aligned} \tag{21}$$

Using the error propagation formula for exponentation,

$$\left(C \pm c_r\right)^2 = C^2 \pm 2\ Cc_r = C^2 \pm 2\ C^2 \hat{c}_r \tag{22}$$

Using the division formula,

$$\frac{AB \pm AB\sqrt{\hat{a}_r^2 + \hat{b}_r^2}}{C^2 \pm 2\,C^2\hat{c}_r} = \frac{AB}{C^2} \pm \sqrt{\frac{A^2B^2\left(\hat{a}_r^2 + \hat{b}_r^2\right)}{C^4} + \frac{A^2B^2C^4 4\;\hat{c}_r^2}{C^8}}$$

$$= \frac{AB}{C^2} \pm \frac{AB}{C^2}\sqrt{\hat{a}_r^2 + \hat{b}_r^2 + 4\;\hat{c}_r^2} \qquad (23)$$

Thus,

$$S_C \pm e_r = \frac{kAB}{C^2D} \pm \frac{kAB}{C^2D}\sqrt{\hat{a}_r^2 + \hat{b}_r^2 + 4\;\hat{c}_r^2} \qquad (24)$$

Therefore.

$$e_r = \frac{kAB}{C^2D}\sqrt{\hat{a}_r^2 + \hat{b}_r^2 + 4\;\hat{c}_r^2} \qquad (25)$$

and

$$\frac{e_r}{kAB/C^2D} = \hat{e}_r = \sqrt{\hat{a}_r^2 + \hat{b}_r^2 + 4\;\hat{c}_r^2} \qquad (26)$$

Propagation of Systematic Uncertainties

In general, the systematic components of uncertainty are unsymmetrical. Therefore, let

$$B_{-b_{s\ell}}^{+b_{su}},\; C_{-c_{s\ell}}^{+c_{su}},\; D_{-d_{s\ell}}^{+d_{su}},\text{ and } S_{C}{}_{-e_{s\ell}}^{+e_{su}}$$

be systematic error bounds on EF, H, F, and S_T, and let $\hat{b}_{su}$, $\hat{b}_{s\ell}$, etc., denote the corresponding relative uncertainties. The upper and lower bounds on γ are taken as m and 1/m, respectively. Since all variables in Equation 16 are non-negative, the upper bound on source severity can be obtained by substituting upper bounds in the numerator and lower bounds in the denominator. Thus,

$$S_C + e_{su} = (kA)\frac{m\left(B + b_{su}\right)}{\left(D - d_{s\ell}\right)\left(C - c_{s\ell}\right)^2} \qquad (27)$$

Rearranging this equation yields

$$\hat{e}_{su} = \frac{m\left(1 + \hat{b}_{su}\right)}{\left(1 - \hat{c}_{s\ell}\right)^2\left(1 - \hat{d}_{s\ell}\right)} - 1 \qquad (28)$$

The lower bound is obtained in a similar manner.

$$\hat{e}_{s\ell} = 1 - \left[\frac{1 - \hat{b}_{s\ell}}{m\left(1 + \hat{d}_{su}\right)\left(1 + \hat{c}_{su}\right)^2}\right] \qquad (29)$$

Uncertainty Interval for Source Severity

The total uncertainty associated with source severity is the sum of the random and systematic components (see Appendix B). Symbolically,

$$\hat{e}_u = \hat{e}_r + \hat{e}_{su}$$

$$\hat{e}_\ell = \hat{e}_r + \hat{e}_{s\ell}$$

Using Equations 26, 28, and 29 for $\hat{e}_r$, $\hat{e}_{su}$, and $\hat{e}_{s\ell}$ yields:

$$\hat{e}_u = \sqrt{\hat{a}_r^{\,2} + \hat{b}_r^{\,2} + 4\,\hat{c}_r^{\,2}} + \frac{m\left(1 + \hat{b}_{su}\right)}{\left(1 - \hat{d}_{s\ell}\right)\left(1 - \hat{c}_{s\ell}\right)^2} - 1 \qquad (30)$$

$$\hat{e}_\ell = \sqrt{\hat{a}_r^{\,2} + \hat{b}_r^{\,2} + 4\,\hat{c}_r^{\,2}} + 1 - \left[\frac{1 - \hat{b}_{s\ell}}{m\left(1 + \hat{d}_{su}\right)\left(1 + \hat{c}_{su}\right)^2}\right] \qquad (31)$$

The uncertainty interval for source severity is then given by the following inequality:

$$S_C\left(1 - \hat{e}_\ell\right) \leq S_T \leq S_C\left(1 + \hat{e}_u\right) \qquad (32)$$

STOCHASTIC DECISION APPROACH

In developing the stochastic decision test approach the true value of source severity is assumed to be a random parameter. The test is developed for a fixed source in a random receiving atmosphere. It treats all uncertainties as random errors. The approach differs from the deterministic decision approach in that the deterministic approach fixes the receiving atmosphere. The approach also differs from the one presented in Appendix H in that the latter uses variable emission sources as well as a random receiving atmosphere.

Uncertainties in the Stochastic Decision Approach

In the deterministic decision approach, S_T is treated as a fixed parameter by treating γ as a constant. It is intuitively reasonable to attempt to describe S_T in a stochastic manner since meteorological change day to day. The calculated source severity is:

$$S_C = \frac{2(CAP)(EF)}{e\mu FH^2}\left(\frac{t_o}{t}\right)^{0.17} \quad (33)$$

which may be rewritten as.

$$S_C = K\,\frac{(CAP)(EF)}{\mu FH^2} \quad (34)$$

where K is a constant. Recalling Equation 16, we have

$$S_T = \gamma\frac{2(CAP)(EF)}{e\mu FH^2}\left(\frac{t_o}{t}\right)^{0.17} \quad (16)$$

which may be rewritten as

$$S_T = \gamma K\,\frac{(CAP)(EF)}{\mu FH^2} \quad (35)$$

Based on the discussions presented earlier, S_T can be treated as a log-normally distributed random parameter. S_T is the product of a constant and a log-normally distributed variable (γ). Treatment of the above equations is best accomplished by taking logarithms.

$$\ln S_C = \ln K = \ln (CAP) + \ln (EF) - \ln \mu - \ln F - 2 \ln H$$

The expected value of S_C ($E(\log S_C)$) is

$$E(\log S_C) = \log\left[K\frac{(CAP)(EF)}{\mu H^2\ F}\right] \quad (36)$$

The variance of $\log S_C$ ($V(\log S_C)$) is

$$V(\log S_C) = \frac{1}{N}\left[S_{CAP}^{\ 2} = 4\ S_A^2 + S_{EP}^{\ 2}\right] \quad (37)$$

Uncertainty Interval for Source Severity

The uncertainty interval for S_T is derived as follows:

$$\ln S_C - Z\alpha\sqrt{V(\log S_C)} \leq \log S_T \leq \ln S_C + Z\alpha\sqrt{V(\log S_C)} \quad (38)$$

TESTS OF HYPOTHESIS

The above discussion leads to the development of two approaches to hypothesis testing for source severity. The deterministic decision approach treats S_T as a fixed parameter and tests whether $S_T \geq 1.0$. The stochastic approach treats S_T as a random parameter; it tests the number of days for which $S_T \geq 1$ againsts an "acceptable" number of such occurances. The following discussion developes the hypothesis testing procedure for both decision approaches.

Hypothesis Testing for Deterministic Decision Approach

Inequality 32 can be used to set up a formal test of hypothesis. The hypothesis that we wish to test (the null hypothesis, H_0) is that the true severity is greater than or equal to unity; i.e.,

$$H_0: \quad S_T \geq 1.0$$

This hypothesis is tested against the alternative hypothesis, H_A.

$$H_A: \quad S_T < 1.0$$

The statistic which is used to make the test is S_C. If S_C exceeds a specified critical value, S_C^*, H_0 is accepted and it is concluded that $S_T \geq 1.0$; if S_C is less than S_C^*, H_0 is rejected and it is concluded that $S_T \leq 1.0$.

The basic idea behind the test is that it should be concluded that $S_T \geq 1.0$ whenever the upper end point of the uncertainty interval (Interval 32) exceeds unity. Thus, the critical value, S_C^*, is given by

$$S_C^* \left(1 + \hat{e}_u\right) = 1.0$$

$$S_C^* = \frac{1}{1 + \hat{e}_u} \tag{39}$$

With any such statistical test, two types of errors can occur. These errors, usually designated Type I and Type II, are defined as follows:

Type I Error: Reject H_0 when it is, in fact, true. That is, we conclude that $S_T < 1.0$ when it is, in fact, greater than unity.

Type II Error: Accept H_0 when it is, in fact, false. That is, we conclude that $S_T \geq 1.0$ when it is, in fact, less than unity.

These possible errors can be represented schematically as follows:

	H_0 is true	H_0 is false
Accept H_0	Correct decision	Type II error
Reject H_0	Type I error	Correct decision

Within the context of the Source Assessment Program, the effects of the two types of incorrect decisions are as follows: a Type I error may results in a needed emissions reduction not being accomplished; a Type II eror may result in an emissions reduction which is not really needed.

In order to see how the risk of an incorrect decision can be controlled, it is convenient to use Inequality 32: as repeated here.

$$S_C\left(1 - \hat{e}_\ell\right) \le S_T \le S_C\left(1 + \hat{e}_u\right) \qquad (32)$$

Now assuming that 95% confidence intervals (i.e., A ± a_r, etc.) were used to construct the approximate Interval (32), and further assuming that the confidence level associated with Interval 32 is greater than or equal to the nominal level, the probability that the true value of S_T will fall in Interval 32 is greater than or equal to 95%. In other words, the probability of obtaining a value of S_T which falls on either side of Interval 32 is less than or equal to 2.5%.

Thus,

$$P\left\{S_C < \frac{S_T}{1 + \hat{e}_u}\right\} \le 2.5\% \qquad (40)$$

Now suppose that the null hypothesis is true, so that $S_T \geq 1.0$. Then we have

$$P\left\{S_C < \frac{1}{1 + \hat{e}_u} \middle| S_T \geq 1.0\right\} \leq P\left\{S_C < \frac{S_T}{1 + \hat{e}_u} \middle| S_T \geq 1.0\right\} \leq 2.5\%$$

Using Equation 39, we obtain:

$$P\left\{S_C < S_C^* \middle| S_T \geq 1.0\right\} \leq 2.5\% \qquad (41)$$

But $S_C < S_C^*$ is precisely the condition for rejecting H_0. Therefore, the probability of rejecting H_0 when it is true, i.e., of making a Type I error, is less than or equal to 2.5%. In general, if we work at the $(1 - 2\alpha)(100\%)$ confidence level in constructing the uncertainty interval (32), then the probability of making a Type I error will be less than or equal to $100\alpha\%$. Thus, by specifying the level of the test, α, the risk of making a Type I error can be made as small as desired.

In a similar manner, when the null hypothesis is false:

$$P\left\{S_C > \frac{1}{1 - \hat{e}_\ell} \middle| S_T < 1.0\right\} < P\left\{S_C > \frac{S_T}{1 - \hat{e}_\ell} \middle| S_T < 1.0\right\} \leq 2.5\%$$

Therefore, if H_0 is accepted when $S_C > 1/(1 - \hat{e}_\ell)$, the probability of a Type II error will be less than 2.5%. However, H_0 is accepted whenever $S_C > S_C^* = 1/(1 + \hat{e}_u)$. Thus, for values of S_C in the interval

$$\frac{1}{1 + \hat{e}_u} \leq S_C \leq \frac{1}{1 - \hat{e}_\ell} \tag{42}$$

there will be a (relatively) large chance of a Type II error.[a] Therefore, in order to minimize the number of Type II errors in repeated applications of the test procedure, the Interval 36 should be made as small as possible; i.e., the smaller the uncertainty in source severity, the fewer the number of incorrect decisions that are expected to be made.

Within the context of the Source Assessment Program, the width of the Interval 36 can be controlled primarily through the random uncertainty in emission factor, $\hat{b}_r$. This uncertainty can (theoretically) be made as small as desired by analyzing a sufficiently large number of samples. The other uncertainties which contribute to Interval 36 cannot readily be controlled in this program.

In order to specify an allowable value for $\hat{b}_r$, a width must be specified for Interval 36. Although this latter specification is arbitrary, an upper bound on $\hat{b}_r$ can be obtained by setting the upper limit in Interval 36 equal to infinity. Roughly speaking, this condition means that we can never be sure that $S_T > 0$, no matter how large the calculated severity, S_C, may be. This can be seen by setting $(1 - \hat{e}_p) = 0$ in Interval 32. Stated another way, when this condition prevails, there will always be a large probability of a Type II error, regardless of how large the value of S_C may be. From Interval 36, this upper bound on $\hat{b}_r$ is found by solving the equation

---

[a] An alternative procedure is to reject H_0 when $S_C < 1/(1 + \hat{e}_u)$, accept H_0 when $S_C > 1/(1 - \hat{e}_\ell)$, and make no decision when S_C falls in the Interval 36. This procedure is considered in more detail below. For the present, it is assumed that a decision (accept or reject H_0) is always required.

$$\hat{e}_\ell = 1.0 \tag{43}$$

Hypothesis Testing for Stochastic Decision Approach

The hypothesis that we wish to test is not whether $S_T \geq 1.0$ but whether the number of times per year $S_T > 1.0$ is more than some acceptable number of times; i.e., H_o: $S_T \geq 1$ on more than r days per year. This hypothesis is tested against the alternative hypothesis, H_A,

$$H_A\text{:}\quad S_T \geq \text{ on less than r days per year}$$

Obviously this hypothesis would be most relevant when the "measure of harmful effect" is proportional to the number of days of threshhold exceedance. If the harm done depends in a more complicated way on S_T, for example on the actual amounts by which S_T exceeds 1.0, then other hypotheses may be more appropriate.

The hypothesis H may be rephrased as:

H_o: The expected number of days on which $S_T > 1$, exceeds r.

or

$$H_o\text{:}\quad E(Z) > r$$

where E denotes the mean (expected value), but, $E(Z) = 365\ P\left\{S_T > 1\right\}$ so that

$$H_o\text{:}\quad P\left\{S_T > 1\right\} > \bar{n}$$

where $\bar{n} = r/365$.

The hypothesis becomes,

$$H_o\text{:}\quad P\left\{\gamma S_C > 1\right\} > \bar{n}$$

$$H_o\text{:}\quad P\left\{\ln\gamma > \ln S_C\right\} > \bar{n}$$

The hypothesis becomes,

$$H_o\text{:}\quad E(\ln S_C) > V(\ln\gamma) Z_{\bar{n}}$$

where V(lnγ) is the variance of log γ (V(lnγ) = 0.3136 from Appendix I) and $Z_{\bar{n}}$ is the $\bar{n}$-percentage point of the standard normal distribution. The level α test for H_0 is to reject H_0 if $\ln S_C < \ln S_C^*$ where

$$P\left\{\ln S_C < \ln S_C^* \middle| \sqrt{V(\ln\gamma)}\, Z_{\bar{n}} = \alpha\right\}$$

so that

$$\frac{\ln S_C^* - \sqrt{V(\ln\gamma)}\, Z_{\bar{n}}}{V(\ln S_C)} = Z_\alpha \qquad (44)$$

or

$$\log S_C^* = (Z_{\bar{n}})\left[V(\log\gamma)\right]^{\frac{1}{2}} + (Z_\alpha)\left[V(\log S_C)\right]^{\frac{1}{2}} \qquad (45)$$

The critical value of S_C for an α-level test is

$$S_C^* = e^{(Z_{\bar{n}})[V(\log\gamma)]^{\frac{1}{2}} + (Z_\alpha)[V(\log S_C)]^{\frac{1}{2}}} \qquad (46)$$

As an example, assume that r = 1 day per year (current basis for several of the criteria pollutants), and that α = 0.05 is desired. We have:

$r = 1$
$\bar{n} = 1/365 = 0.00274$
$Z_{\bar{n}} = -2.786$
$Z_\alpha = Z_{0.05} = -1.645$

Hence, we would reject H_0 if,

$$S_C < e^{-\left(2.776\sqrt{V(\log\gamma)} + 1.645\sqrt{V(\log S_C)}\right)} \qquad (47)$$

Remembering that

$$V(\log\gamma) = 0.3134$$

or

$$[V(\log\gamma)]^{\frac{1}{2}} = 0.560$$

we would reject H_0 if

$$S_C < e^{-\left(1.555 + 1.645\, V(\log S_C)\right)} \qquad (48)$$

COMPARISON OF ALTERNATE DECISION APPROACHES

The above approaches to hypothesis testing utilize different starting assumptions (S_T is a "fixed" versus a "random" parameter) and they treat uncertainties in different manners (large uncertainties treated as systematic errors versus all uncertainties treated as random errors). Yet, the critical values of source severity for decision making are quite similar. This is best illustrated in Figure 2 where S_c^* is plotted as a function of the alternative decision appraches and uncertainty in emission factor. The most frequently occuring range of $\hat{b}_r$ (emission factor uncertainties) is 0.6 to 1.0. The bias (b_{su}) in available sampling and analysis technologies ranges from 0.0 to 0.5. The values S_c^* vary from 0.5 to 0.8 for the general cases displayed in Figure 2 regardless of the decision approach used.

The following observations of Figure 2 are readily apparent. Given the current (1979) capabilities in measurement technologies, both decision approaches give the same S_c^* values. The deterministic decision approach is insensitive to improvements in measurement technologies. The stochastic decision approach is quite sensitive to measurement technology improvements.

Figure 3 shows how S_c^* changes as r increases for the stochastic decision approach.

GUIDELINES FOR SOURCE ASSESSMENT PROGRAM

The above test of hypothesis can be applied to the decisionmaking process by computing the critical severity, S_c^*, for each pollutant emitted from each source type; i.e., the test can be performed for each individual pollutant and each source type studied. An example of this approach is given in Appendix J. Alternatively, the test procedure can be used to derive general guidelines for decision making in the Source Assessment Program. The latter approach is followed in this section.

In order to establish overall guidelines for the Source Assessment Program, generalized error bounds have been estimated as shown in Table 3. These estimates are derived in Appendix I. The values listed for $\hat{a}_r$ and $\hat{c}_r$ represent 95% confidence limits. In this analysis, $\hat{b}_r$ is considered to be an independent variable. The values given for $\hat{b}_r$ in Table 3 indicate the range in which b_r can be expected to fall in most cases. Three sets of values are listed for the uncertainty in "acceptable" concentration, F. In Case A, there is no uncertainty associated with F (cf. discussion on page 30). In Case B_1, the "acceptable" concentration is considered to be uncertain, and the TLV conversion factor, G, is equal to the mean value (0.047) for criteria pollutants (see Appendices E and I). Case B_2 is the same as B_1 except that G = 0.0033, the value currently used in the Source Assessment Program.

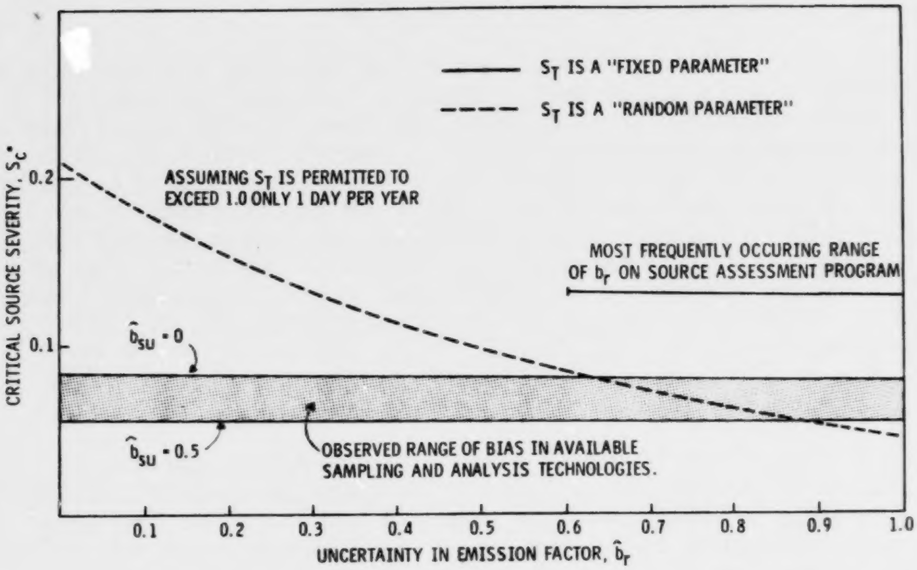


Figure 2. Impact of alternative hypothesis test approaches on S_C^*

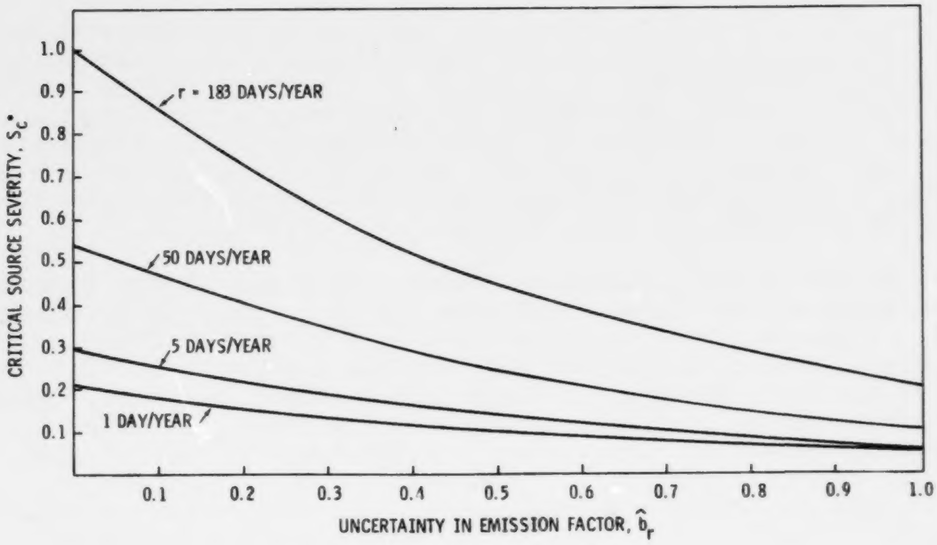


Figure 3. Variation in critical source severity as a function of emission factor data quality and the acceptable number of days for severity to exceed 1.0 when treating the true source severity as a random parameter.

TABLE 3. GENERALIZED ERROR BOUNDS FOR SOURCE ASSESSMENT

Uncertainty	Value		
$\hat{a}_r$	0.05		
$\hat{c}_r$	0.05		
$\hat{b}_r$	0 to 1.0		
$\hat{b}_{su}$	0.5		
$\hat{b}_{s\ell}$	0.1		
$\hat{c}_{su}$	0.1		
$\hat{c}_{s\ell}$	0.5		
$\hat{d}_{su}$	A 0	B_1 4.2	B_2 71.6
$\hat{d}_{s\ell}$	A 0	B_1 0.81	B_2 0
m	3.0		

For the Deterministic Decision Approach

Substituting values from Table 3 into Equations 30 and 31 yields the following error bounds for source severity:

Case A

$$\left.\begin{aligned} \hat{e}_u &= \sqrt{\hat{b}_r^{\,2} + 0.01} + 17 \\ \hat{e}_\ell &= \sqrt{\hat{b}_r^{\,2} + 0.01} + 0.75 \end{aligned}\right\} \qquad (49)$$

Case B_1

$$\left.\begin{aligned} \hat{e}_u &= \sqrt{\hat{b}_r^{\,2} + 0.01} + 94 \\ \hat{e}_\ell &= \sqrt{\hat{b}_r^{\,2} + 0.01} + 0.95 \end{aligned}\right\} \qquad (50)$$

Case B_2

$$\left.\begin{aligned}\hat{e}_u &= \sqrt{\hat{b}_r^2 + 0.01} + 17 \\ \hat{e}_\ell &= \sqrt{\hat{b}_r^2 + 0.01} + 1.0\end{aligned}\right\} \quad (51)$$

The corresponding uncertainty interval for source severity is obtained by substituting Equation 49, 50, or 51 into Inequality 32. For example, taking $b_r = 1.0$ as a worst case,[a]

Case A $$0 \leq S_T \leq 19\ S_C \quad (52)$$

Case B_1 $$0 \leq S_T \leq 96\ S_C \quad (53)$$

Case B_2 $$0 \leq S_T \leq 19\ S_C \quad (54)$$

At the other extreme, when $\hat{b}_r = 0$,

Case A $$0.15\ S_C \leq S_T \leq 18\ S_C \quad (55)$$

Case B_1 $$0 \leq S_T \leq 95\ S_C \quad (56)$$

Case B_2 $$0 \leq S_T \leq 18\ S_C \quad (57)$$

The critical value of the "calculated" severity is given by Equation 39. Substituting for $\hat{e}_u$ from Equations 49, 50, and 51 yields the following results:

Case A $$S_C^* = \frac{1}{18 + \sqrt{\hat{b}_r^2 + 0.01}} \quad (58)$$

Case B_1 $$S_C^* = \frac{1}{95 + \sqrt{\hat{b}_r^2 + 0.01}} \quad (59)$$

Case B_2 $$S_C^* = \frac{1}{18 + \sqrt{\hat{b}_r^2 + 0.01}} \quad (60)$$

---

[a] Note that source severity cannot be negative, so the lower bound cannot be less than zero.

As $\hat{b}_r$ varies from 0 to 1.0, S_C^* varies from 0.055 to 0.053 in Cases A and B_2 and from 0.0105 to 0.0104 in Case B_1. It follows that the critical test value should be set at 0.05 for Cases A and B_2 and at 0.01 for Case B_1.

For the Stochastic Decision Approach

Substituting values from Table 3 into Equation 26 yields the following value for $V(\log S_C)$, regardless of how F is treated.

$$V(\log S_C) = \sqrt{\hat{b}_r^2 + 0.01} \qquad (61)$$

The corresponding uncertainty interval for source severity is obtained by substituting Equation 61 into Equation 48. For example, taking $\hat{b}_r = 1.0$ as a worst case, we get,

$$0.0404\ S_C \le S_T \le 24.7\ S_C \qquad (62)$$

The corresponding uncertainty interval for source severity is obtained by substituting by taking $\hat{b}_r = 0.0$.

$$0.179\ S_C \le S_T \le 5.58\ S_C \qquad (63)$$

ALLOWABLE RANDOM UNCERTAINTY IN EMISSION FACTOR

An upper bound on the allowable value of $\hat{b}_r$ is obtained by solving Equation 44. Substituting for $\hat{e}_\ell$ from Equations 49, 50, and 51 gives the following results:

	Deterministic Approach		Stochastic Approach
Case A	$\hat{b}_r < 0.23$	(59)	infinity
Case B_1	No solution		
Case B_2	No solution		

In Cases B_1 and B_2, the restriction imposed by Equation 44 cannot be met even with $\hat{b}_r = 0$. Hence, in these cases, there will always be a large risk of a Type II error in the screening procedure.

Strictly speaking, Equations 58 through 61 are valid for a test level, α, of 0.025. However, these results are nearly independent of α. This is a reflection of the fact that the total uncertainty in source severity is dominated by the systematic component of uncertainty (see below).

In principle, then, for Case A (no uncertainty in "acceptable" concentration), Inequality 61 can be used with standard procedures to estimate the number of samples required in a sampling program. In most situations involving environmental sampling, however, the restriction imposed by Equation 61 will result in an impractically large number of samples. In practice then, there is generally little that can be done to control the risk of a Type II error when using the present screening procedures. This is again a reflection of the fact that the random uncertainty in emission factor makes a relatively small contribution to the total uncertainty in source severity.

EFFECT OF RANDOM UNCERTAINTY IN EMISSION FACTOR

Comparison of Inequalities 52 through 53 with Inequalities 55 through 57 shows that the random uncertainty in emission factor has a relatively small effect on the total uncertainty in source severity. The total uncertainty is dominated by the systematic uncertainties, as illustrated in Table 4. This table lists the reduction in the width of the uncertainty interval (Inequality 41, 42, or 43) which would be obtained if each of the individual uncertainties was eliminated while leaving the others unchanged. The table was constructed using the values of the individual uncertainties listed in Table 3, together with $\hat{b}_r = 1.0$, to calculate the width $(\hat{e}_u + \hat{e}_\ell)$ of the Interval 32. Each of the individual uncertainties was then set to zero while leaving the others unchanged, and the resulting reduction in the width of the interval was computed.

TABLE 4. CONTRIBUTION OF INDIVIDUAL UNCERTAINTIES TO TOTAL UNCERTAINTY IN SOURCE SEVERITY (DETERMINISTIC DECISION APPROACH)

Uncertainty eliminated	Reduction in width of uncertainty interval for source severity, %[a]		
	Case A[b]	Case B_1[c]	Case B_2[d]
1. Random uncertainty in emission factor.	10	2	10
2. Systematic uncertainty in emission factor.	30	33	30
3. Total uncertainty in emission factor (1 + 2).	40	35	40
4. Uncertainty in dispersion equation.	63	65	60
5. Uncertainty in plume rise.	69	74	68
6. Total uncertainty in dispersion modeling (4 + 5).	87	90	83
7. Uncertainty in "acceptable" concentration.	0	80	1.3

[a] Uncertainty interval computed assuming $\hat{b}_r = 1.0$.

[b] No uncertainty in acceptable concentration.

[c] Acceptable concentration uncertain; TLV conversion factor equal to geometric mean value.

[d] Acceptable concentration uncertain; TLV conversion factor equal to 1/300.

Contrary to what might be expected, Case B_2, in which the acceptable concentration is "intentionally underestimated," has a narrower uncertainty interval than does Case B_1. In general, underestimating the acceptable concentration decreases both the upper bound (which tends to decrease the total uncertainty) and the lower bound (which tends to increase total uncertainty) on source severity. However, in Case B_1 the lower bound is already zero, so it cannot be decreased further. Therefore, the only effect of underestimating the acceptable concentration is to decrease the upper bound, which reduces the total uncertainty in source severity.

In addition, since $\hat{e}_u$ can vary from zero to infinity while $\hat{e}_\ell$ can vary from zero to one, the above calculation tends to give more weight to the upper bound than to the lower bound. Although underestimating the "acceptable" concentration does not reduce the uncertainty associated with acceptable concentration, it does shift it to the lower bound where it is less conspicuous. Therefore, in Table 4, Case B_1 is more indicative of the relative magnitude of the uncertainty associated with acceptable concentration than is Case B_2.

The upper and lower error bounds on source severity are tabulated in Table 5 for each of the cases listed in Table 4. These error bounds correspond to the situation in which $a_r = c_r = 0$, so that the random component of uncertainty becomes

$$\hat{e}_r = \sqrt{\hat{a}_r^{\,2} + \hat{b}_r^{\,2} + 4\,\hat{c}_r^{\,2}} = \hat{b}_r \tag{64}$$

TABLE 5. CONTRIBUTION OF INDIVIDUAL UNCERTAINTIES TO ERROR BOUNDS ON SOURCE SEVERITY[a] (DETERMINISTIC DECISION APPROACH)

	Lower bound, $1 - \hat{e}_\ell$			Upper bound, $1 + \hat{e}_u$		
Uncertainty eliminated	Case A[b]	Case B_1[c]	Case B_2[d,e]	Case A[b]	Case B_1[c]	Case B_2[d]
1. None.	$0.25 - \hat{b}_r$	$0.05 - \hat{b}_r$	0.00	$18 + \hat{b}_r$	$95 + \hat{b}_r$	$18 + \hat{b}_r$
2. Random uncertainty in emission factor.	0.25	0.05	0.00	18	95	18
3. Systematic uncertainty in emission factor.	$0.28 - \hat{b}_r$	$0.05 - \hat{b}_r$	0.00	$12 + \hat{b}_r$	$63 + \hat{b}_r$	$12 + \hat{b}_r$
4. Total uncertainty in emission factor (2 + 3).	0.28	0.05	0.00	12	63	12
5. Uncertainty in dispersion equation.	$0.74 - \hat{b}_r$	$0.14 - \hat{b}_r$	$0.01 - \hat{b}_r$	$6 + \hat{b}_r$	$32 + \hat{b}_r$	$6 + \hat{b}_r$
6. Uncertainty in plume rise.	$0.30 - \hat{b}_r$	$0.06 - \hat{b}_r$	0.00	$4.5 + \hat{b}_r$	$24 + \hat{b}_r$	$4.5 + \hat{b}_r$
7. Total uncertainty in dispersion modeling (5 + 6).	$0.90 - \hat{b}_r$	$0.17 - \hat{b}_r$	$0.01 - \hat{b}_r$	$1.5 + \hat{b}_r$	$7.9 + \hat{b}_r$	$1.5 + \hat{b}_r$
8. Uncertainty in "acceptable" concentration.	$0.25 - \hat{b}_r$	$0.25 - \hat{b}_r$	$0.25 - \hat{b}_r$	$18 + \hat{b}_r$	$18 + \hat{b}_r$	$18 + \hat{b}_r$

[a] $S_C(1 - \hat{e}_\ell) \leq S_T \leq S_C(1 + \hat{e}_u)$

[b] No uncertainty in acceptable concentration.

[c] Acceptable concentration uncertain; TLV conversion factor equal to geometric mean value.

[d] Acceptable concentration uncertain; TLV conversion factor equal to 1/300.

[e] Since source severity is nonnegative, the lower bound cannot be less than zero.

From Table 5, it can be seen that in Case B_2 the uncertainty associated with acceptable concentration has no effect on the upper bound since the values for Case B_2 are identical to those for Case A (no uncertainty in acceptable concentration). On the other hand, the lower bound is dominated by the uncertainty in acceptable concentration in Case B_2. In fact, the lower bound is essentially zero for Case B_2 in all instances except Item 8, which corresponds to no uncertainty in acceptable concentration.

Table 5 also shows how each of the individual uncertainties contributes to the departure from the ideal situation in which both the upper and lower bounds are equal to unity. (In this situation, Inequality 32 becomes $S_C \leq S_T \leq S_C$; i.e., $S_T = S_C$.) The ideal situation is approached only in Item 7, Case A (no uncertainty in dispersion modeling or acceptable concentration) and then only when the random uncertainty in emission factor, $\hat{b}_r$, is small.

OPERATING CHARACTERISTICS OF THE TEST

In order to fully describe the risk of a Type II error in a test of hypothesis such as that described in this section, it is necessary to construct an operating characteristic curve for the test. The probability of a Type II error can be expressed as a function either of the "true" value (S_T in this case) or the estimated value (S_C) of the parameter in question.

In the Neyman-Pearson theory of statistical inference, the true value of the unknown parameter, S_T, is regarded as a fixed entity, and the measurements, S_C, are visualized as distributed about this value. This situation is illustrated in Figure 4.

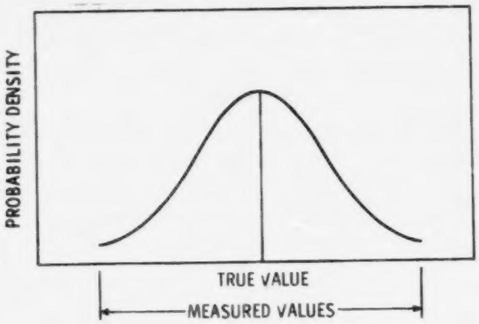


Figure 4. Schematic representation of the fiducial statistical approach.

Deterministic Decision Approach

Operating characteristic (OC) curves are given in Figure 5 for Case A (no uncertainty in "acceptable" concentration) with $\hat{b}_r = 0.1$.[a] The curves give the probability that $S_T \leq 1.0$ as a function of the true severity, S_T. For $S_T < S_C(1 - \hat{e}_\ell)$, the curves give the probability of not making a Type I error; for $S_T \geq S_C(1 + \hat{e}_u)$, they give the probability of making a Type II error. A family of OC curves is obtained with the systematic error in source severity as a parameter. In Figure 5, three members of this family of curves are shown corresponding to no systematic error, systematic error equal to e_{su}, and systematic error equal to $e_{s\ell}$. The latter two curves represent the extreme cases; i.e., all OC curves for this problem lie between these two curves.

---

[a] Construction of these curves is discussed in Appendix F.

Note that in Figure 5 the left-most curve (systematic error equal to e_{su}) represents the worst case with respect to Type I errors, while the right-most curve (systematic error equal to $e_{s\ell}$) represents the worst case with respect to Type II errors. Hence, a convenient means of representing the family of OC curves is a combination of the upper half of the left-most curve and the lower half of the right-most curve, as shown in Figure 6. (The other halves of the curves are of lesser interest since they correspond to error probabilities of 50% or greater.) The curve in Figure 6 is termed an "effective" OC curve since it corresponds to treating the Interval 32 as an exact 95% confidence interval.

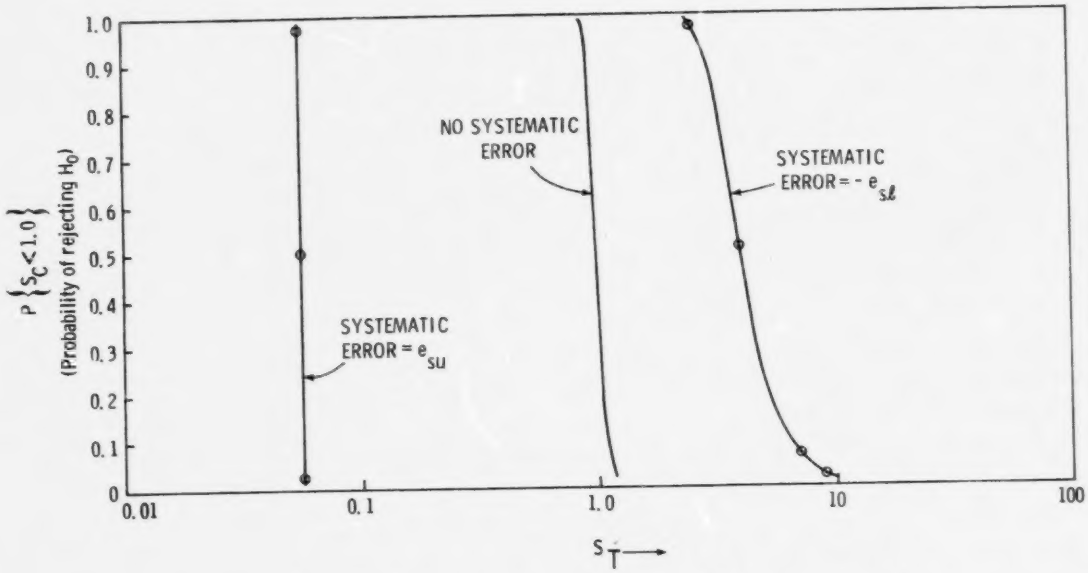


Figure 5. Operating characteristic curves for Case A with $\hat{b}_r$ = 0.10.

The 0.975 and 0.025 probability points on the curve in Figure 7 represent the end points of the uncertainty interval for source severity. The horizontal portion of the curve represents the portion of the uncertainty interval which is due to the systematic component of uncertainty. The left- and right-hand branches of the curve represent the contribution of the random component of uncertainty. The right-hand branch of the curve gives the worst case probability of a Type II error. The worst case probability is greater than or equal to the value given by the horizontal and left-hand branches of the curve. (It is of little consequence whether the worst case probability of a Type II error is, for example, 0.5 or 0.9. The important point is that there is a large risk of a Type II error in this region, and this is clearly indicated by the "effective" curve.) Thus, the graph shows that the worst case probability of a Type II error cannot be reduced to less than 0.5 in this region by reducing the random uncertainty alone. This result can only be achieved by reducing the systematic component of uncertainty.

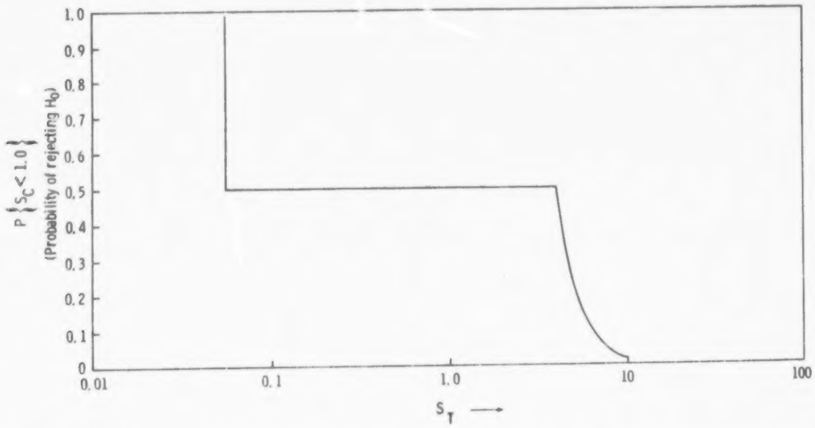


Figure 6. "Effective" operating characteristic curve for Case A with $\hat{b}_r = 0.10$.

In Figure 7, "effective" OC curves are given for Case A (no uncertainty in "acceptable" concentration) as a function of $\hat{b}_r$, the random uncertainty in emission factor.[a] (The curves for Cases B_1 and B_2 are similar except that the horizontal portions of the curves extend to infinity for all values of $\hat{b}_r$.) The curves show that when $\hat{b}_r \geq 0.23$, the worst case probability of a Type II error is greater than or equal to 0.5 for all values of the estimated severity, S_C (cf. Equation 48, Inequality 59, and the related discussions).

---

[a] The left-hand branch of the curve is insensitive to the value of $\hat{b}_r$ due to the dominant effect of the systematic component of uncertainty.

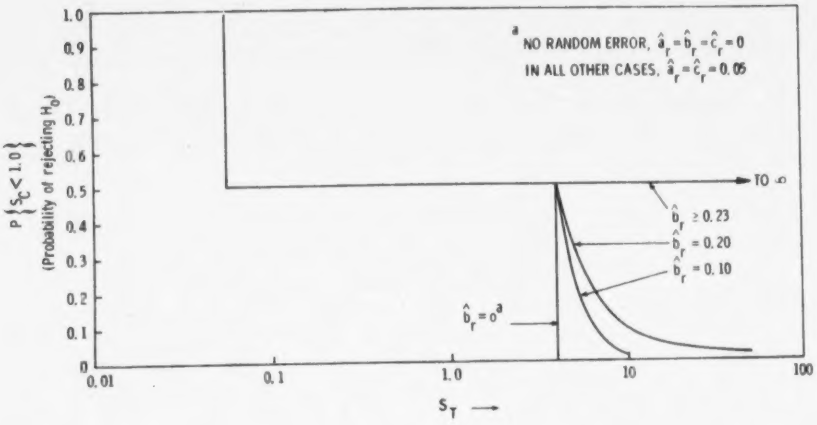


Figure 7. "Effective operating characteristic curves for Case A as a function of random uncertainty in emission factor.

Stochastic Decision Approach

The operating characteristic curve for the stochastic decision approach is given in Figure 7. OC curves for r = 1 and r = 183 are shown; corresponding curves for other values of r are parallel to those shown in Figure 8. Since the stochastic decision test approach treats all uncertainties as random errors, there is no plateau in the curves as found in the previous OC curves.

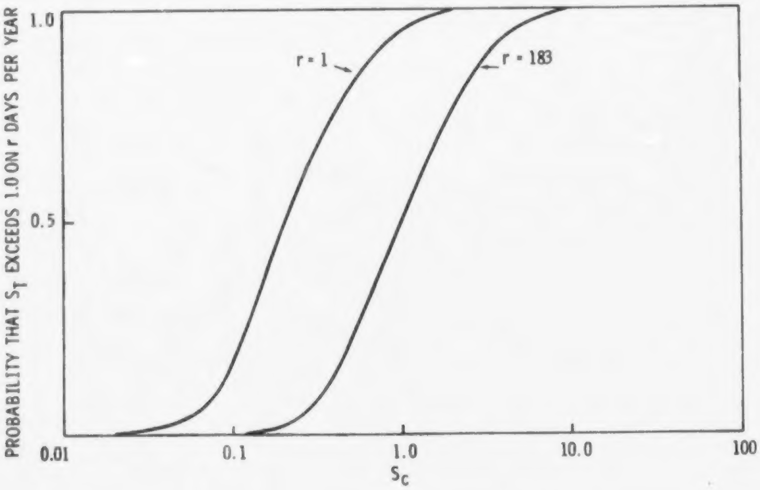


Figure 8. Operating characteristic curve for the stochastic decision test approach.

SUMMARY

The results of the above analysis can be summarized as follows:

- The stochastic decision test approach should be used in preference to the deterministic decision test approach.
- The critical value, S_c^*, for the calculated source severity should be set at 0.05 to 0.08.
- In general, there is no justification for taking a large number of samples to obtain a very precise estimate for the emission factor, since precision in source severity is limited by the precision in available measurement technologies.

The last conclusion should not be construed as implying that sampling and analytical procedures are unimportant in the Source Assessment Program. In the first place, bias in sampling and analysis does have a significant effect on the total uncertainty as shown in Table 4. Secondly, in many cases the effect of random uncertainty in emission factor will be considerably greater than indicated in Table 4. For sources having negligible plume rise, for example, the total uncertainty will be reduced by 69% to 74% over that used in Table 4, and the contribution from the random uncertainty in emission factor will be correspondingly greater. Thirdly, reliable emissions data are of value in their own right, above and beyond their utility in estimating source severity.

SECTION 6

SOURCE ASSESSMENT: EMISSIONS BURDEN

INTRODUCTION

For criteria pollutants, the source assessment screening process makes use of a second decision index, the national emissions burden, in addition to source severity. For a given source type and given criteria pollutant, the national emissions burden, N_B, is defined as follows:

$$N_B = \frac{M_p}{M_n} \qquad (65)$$

where M_p = annual mass emissions of given criteria pollutant from the given source type, kg

M_n = annual mass emissions of given pollutant from all stationary sources nationwide, kg

Equation 65 defines the "true" value of N_B. In practice, it is calculated as follows:

$$N_B = \frac{\left(CAP_T\right)\left(EF_R\right)}{M_{NEDS}} \qquad (66)$$

where CAP_T = total production capacity of source type, kg/yr

EF_R = representative emission factor for source type, g/kg

M_{NEDS} = estimate of M_n obtained from 1972 National Emissions Report (11),[a] kg

---

[a] The National Emissions Report contains data on emissions from both stationary and mobile sources. Only the data for stationary sources are used in estimating M_n.

----

(11) 1972 National Emissions Report; National Emissions Data System (NEDS) of the Aerometric and Emissions Reporting System (AEROS). EPA-450/2-74-012, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, June 1974. 422 pp.

Note that the emission factor in Equation 53 is not the same as that used to calculate source severity. The latter emission factor is for a single emission point within a representative plant, while the emission factor in Equation 53 is the total for all emission points within a representative plant.

As noted in Section 5, the national emissions burden is presently used for screening a source type as follows: the source type is considered to be a potential candidate for control technology development only if $N_B \geq 0.001$.

The objective of this section is to perform an analysis for N_B similar to that given in the previous section for source severity.

UNCERTAINTY IN N_B

The uncertainty in N_B arises from the following components:

- There is a random uncertainty in production capacity as discussed in the previous section.
- There is uncertainty in the emission factor which has both random and systematic components.
- There is uncertainty in nationwide annual emissions, due to the use of the NEDS data base, which has both random and systematic components.

For convenience, let A, B, and D denote the "measured" values of the variables as follows:

$$A = CAP_T$$
$$B = EF_R$$
$$D = M_n$$

Let $\hat{a}_r$, $\hat{b}_r$, and $\hat{d}_r$ be the relative random uncertainties associated with A, B, and D. Using the error propagation formulas from Table B-1, the random uncertainty in N_B is found to be:

$$\hat{e}_r = \sqrt{\hat{a}_r^{\,2} + \hat{b}_r^{\,2} + \hat{d}_r^{\,2}} \tag{67}$$

Considering next the systematic uncertainty, let

$$B\,{}^{+b_{su}}_{-b_{s\ell}} \quad \text{and} \quad D\,{}^{+d_{su}}_{-d_{s\ell}}$$

be systematic error bounds on EF_R and M_n, and let $\hat{b}_{su}$, $\hat{b}_{s\ell}$, etc., denote the corresponding relative error bounds. Using the formulas from Table B-8 yields the following systematic error bounds on N_B:

$$\hat{e}_{su} = \frac{\hat{b}_{su} + \hat{d}_{s\ell}}{1 + \hat{d}_{s\ell}} \tag{68}$$

$$\hat{e}_{s\ell} = \frac{\hat{b}_{s\ell} + \hat{d}_{su}}{1 + \hat{d}_{su}} \tag{69}$$

The total uncertainty in N_B is the sum of the random and systematic components:

$$\hat{e}_u = \sqrt{\hat{a}_r^2 + \hat{b}_r^2 + \hat{d}_r^2} + \frac{\hat{b}_{su} + \hat{d}_{s\ell}}{1 - \hat{d}_{s\ell}} \tag{70}$$

$$\hat{e}_\ell = \sqrt{\hat{a}_r^2 + \hat{b}_r^2 + \hat{d}_r^2} + \frac{\hat{b}_{s\ell} + \hat{d}_{su}}{1 - \hat{d}_{su}} \tag{71}$$

Thus, we have the following uncertainty interval for the true value of N_B:

$$(N_B)_C\ (1 - \hat{e}_\ell) \leq (N_B)_T \leq (N_B)_C\ (1 + \hat{e}_u) \tag{72}$$

where $\hat{e}_u$ and $\hat{e}_\ell$ are given by Equations 70 and 71 and $(N_B)_C$ is the estimated or "calculated" value of N_B; i.e.,

$$(N_B)_C = \frac{A(B)}{D} \tag{73}$$

For the purpose of this example, the following error bounds will be used:

$\hat{a}_r = 0.05$

$\hat{d}_r = 0.01$

$\hat{b}_{su} = 0.5$

$\hat{b}_{s\ell} = 0.1$

$\hat{d}_{su} = 0.3$

$\hat{d}_{s\ell} = 0$

The values for $\hat{a}_r$, $\hat{b}_{su}$, and $\hat{b}_{s\ell}$ are the same as those used in the previous section. The value for $\hat{d}_r$ was obtained as follows: the variability in each category in the 1972 National Emissions Report (11) is estimated to be ±5% (personal communication with J. Bosch, U.S. Environmental Protection Agency, Washington, D.C., 1 December 1976). There are approximately 100 categories which contribute to the national totals. Hence, the variability in the totals is estimated (assuming all categories contribute equally to the total) to be:

$$\frac{\sqrt{100(0.05)^2}}{100} = \pm 0.005$$

The 95% confidence limits are approximately twice this value, or ±0.01.

The systematic error bounds on M_n were obtained by assuming that the NEDS totals are systematically low due to incomplete inventories. Assuming the emissions inventories are at least 70% complete gives $\hat{d}_{s\ell} = 0$ and $\hat{d}_{su} = 0.3$.

With the above values, Equations 70 and 71 become:[a]

$$\hat{e}_u \cong \hat{b}_r + 0.5$$

$$\hat{e}_\ell \cong \hat{b}_r + 0.3$$

The uncertainty interval for N_B is then:

$$(N_B)_C \; (0.7 - \hat{b}_r) \leq (N_B)_T \leq (N_B)_C \; (\hat{b}_r + 1.5) \qquad (74)$$

STATISTICAL TEST OF HYPOTHESIS

Inequality 74 can be used as in the previous section to formally set up a statistical test of hypothesis. The null hypothesis in this case is:

$$H_0: \quad (N_B)_T \geq 0.001$$

This hypothesis will be accepted if the upper limit in Inequality 61 exceeds 0.001. Therefore, the critical value for the test, $(N_B^*)_C$, is given by:

$$(N_B^*)_C = \frac{0.001}{\hat{b}_r + 1.5} \qquad (75)$$

Inverting Inequality 74 yields:

$$\frac{(N_B)_T}{\hat{b}_r + 1.5} \leq (N_B)_C \leq \frac{(N_B)_T}{0.7 - \hat{b}_r} \qquad (76)$$

The upper end point of this interval becomes infinite for $\hat{b}_r = 0.7$, which places an upper bound on the acceptable uncertainty in emission factor. With this value of $\hat{b}_r$, Equation 62 becomes $(N_B^*)_C \cong 0.0005$.

As $\hat{b}_r$ ranges between zero and one, $(N_B^*)_C$ varies from 0.00067 to 0.00042. Since the level of the test, α, enters the calculation only through $\hat{b}_r$, the critical test value is again nearly independent of the test level.[a]

SUMMARY

The above analysis leads to the following guidelines for use of the second source assessment decision index:

1. The critical value of national emissions burden should be set at 0.0005 = 0.05%.
2. The random uncertainty in emission factor should be less than ±70%.

The emission factor referred to here is the total emission factor for the representative plant. However, the above restriction can be met if the uncertainty in the emission factor for each emission point in the representative plant is required to be less than ±70%.

The analysis of the first source assessment decision index resulted in an upper bound on the random uncertainty in emission factor for criteria pollutants of ±23%.[b] Hence, the present value of ±70% theoretically places no additional restriction on sampling and analytical procedures. In practice, however, just the opposite may be true. The value ±23% may be impractical and, therefore, may be ignored entirely, while the value ±70% may be reasonable, and, therefore, represent a real restriction

---

[a] $\hat{a}_r$ has been neglected compared to $\hat{b}_r$.

[b] Criteria pollutants are covered by Case A.

SECTION 7

OFFSET CALCULATIONS: PLANT EXPANSION PROBLEM

INTRODUCTION

As part of EPA's overall program to achieve and maintain National Ambient Air Quality Standards, the individual states are required (through State Implementation Plans) to determine the amount of emissions reduction necessary to offset the probable impact of increased population, industrial activity, motor vehicle traffic, and other growth factors. Of particular concern in this section is the policy regarding increased emissions due to increased industrial activity through expansions of existing plants.

One way to offset emissions from a plant expansion is to require the company to make a corresponding reduction in emissions from its existing unit. Such a policy entails the associated problem of determining *a posteriori* whether the company is complying with the policy. Such a determination is not entirely straightforward because the emission rates are not known exactly. This uncertainty should be taken into account in the decision-analysis process.

The objective of this section is to analyze the uncertainty associated with the above problem and construct an approximate statistical test as the basis for making a decision regarding compliance with the offset policy.

WORKING EQUATIONS

The difference in emission rates of a given pollutant before and after plant expansion can be expressed as follows:

$$Q_2 = Q_1 - \Delta Q_1 + Q_N \qquad (77)$$

or

$$Q_2 - Q_1 = Q_N - \Delta Q_1$$

where Q_2 = emission rate after expansion, g/s
Q_1 = original emission rate, g/s
Q_N = emission rate from new unit, g/s
ΔQ_1 = reduction in emission rate from original unit, g/s

Both Q_N and ΔQ_1 are composed of terms corresponding to the individual emission points within the plant. Thus,

$$Q_N = \sum_i Q_{Ni} \quad \text{(sum over all emission points in new unit)} \tag{78}$$

and

$$\Delta Q_1 = \sum_j (\Delta Q_1)_j \quad \text{(sum over all emission points in original unit which have different emissions after expansion)}$$

$$= \sum_j (Q_{1j} - Q_{1'j})$$

$$\Delta Q_1 = \sum_j Q_{1j} - \sum_j Q_{1'j} \tag{79}$$

In these equations Q_{1j} and $Q_{1'j}$ denote emission rates from the j^{th} emission point in the original unit before and after expansion, respectively.

Equations 77, 78, and 79 are the basic relationships governing the problem. Note that the difference $(Q_2 - Q_1)$ depends only on emissions from the new unit and those emissions from the original unit which are changed in the expansion.

UNCERTAINTY IN $(Q_2 - Q_1)$

Uncertainty in $(Q_2 - Q_1)$ is due to uncertainties in the individual emission rates Q_{Ni}, Q_{1j}, and $Q_{1'j}$. Once the random and systematic uncertainties in these values have been determined, the uncertainty in $(Q_2 - Q_1)$ can be obtained using error propagation formulas for addition and subtraction as outlined below.

Consider first the random uncertainty. Let B_N and B_Δ denote the measured values of Q_N and ΔQ_1, and let b_{rn} and $b_{r\Delta}$ be the corresponding random uncertainties. Then the random uncertainty in $(Q_2 - Q_1)$ is, according to the subtraction formula in Table B-1, Appendix B:

$$e_r = \sqrt{b^2_{rN} + b^2_{r\Delta}}$$

where

$$b^2_{rN} = \sum_i b^2_{rNi}$$

and

$$b^2_{r\Delta} = \sum_j b^2_{r1j} + \sum_j b^2_{r1'j} = \sum_j \left(b^2_{r1j} + b^2_{r1'j}\right)$$

Therefore, we have:

$$e_r = \left[\sum_i b^2_{rNi} + \sum_j \left(b^2_{r1j} + b^2_{r1'j}\right)\right]^{1/2} \qquad (80)$$

where b_{rNi}, b_{r1j}, and $b_{r1'j}$ are the random uncertainties associated with Q_{Ni}, Q_{1j}, and $Q_{1'j}$, respectively.

Considering next the systematic uncertainty, let

$$(B_N)^{+b_{uN}}_{-b_{\ell N}}, \quad (B_\Delta)^{+b_u}_{-b_{\ell\Delta}}, \quad \text{and } (B_N - B_\Delta)^{+e_{su}}_{-e_{s\ell}}$$

be systematic error bounds for Q_N, ΔQ_1, and $(Q_2 - Q_1)$. Then, using the addition and subtraction formulas from Table B-8, we obtain:

$$e_{su} = b_{uN} + b_{\ell\Delta} = \sum_i b_{uNi} + \sum_j \left(b_{\ell 1j} + b_{u1'j}\right) \qquad (81)$$

$$e_{s\ell} = b_{\ell N} + b_{u\Delta} = \sum_i b_{\ell Ni} + \sum_j \left(b_{u1j} + b_{\ell 1'j}\right) \qquad (82)$$

where b_{uNi}, $b_{\ell 1j}$, $b_{u1'j}$, etc., are systematic bounds corresponding to Q_{Ni}, Q_{1j} and $Q_{1'j}$.

The total uncertainty in $(Q_2 - Q_1)$ is the sum of the random and systematic components. Combining Equations 80, 81, and 82, we obtain:

$$e_u = \left[\sum_i b^2_{rNi} + \sum_j \left(b^2_{r1j} + b^2_{r1'j}\right)\right]^{1/2} + \sum_i b_{uNi} + \sum_j \left(b_{\ell 1j} + b_{u1'j}\right) \qquad (83)$$

$$e_\ell = \left[\sum_i b^2_{rNi} + \sum_j \left(b^2_{r1j} + b^2_{r1'j}\right)\right]^{1/2} + \sum_i b_{\ell Ni} + \sum_j \left(b_{u1j} + b_{\ell 1'j}\right) \qquad (84)$$

The uncertainty interval for $(Q_2 - Q_1)$ is given by:

$$B_N - B_\Delta - e_\ell \leq (Q_2 - Q_1) \leq B_N - B_\Delta + e_u \qquad (85)$$

Alternatively, Inequality 72 may be written as:

$$(Q_2 - Q_1)_C - e_\ell \leq (Q_2 - Q_1)_T \leq (Q_2 - Q_1)_C + e_u \qquad (86)$$

where $(Q_2 - Q_1)_T$ denotes the true value of the emission rate difference and $(Q_2 - Q_1)_C \equiv B_N - B_\Delta$ is the calculated difference.

TEST OF HYPOTHESIS

Inequality 86 forms the basis for a formal test of hypothesis. The hypothesis to be tested is:

$$H_0: \ (Q_2 - Q_1)_T > 0; \text{ i.e., } Q_2 > Q_1$$

This is tested against the alternative hypothesis:

$$H_A: \ (Q_2 - Q_1)_T \leq 0; \text{ i.e., } Q_2 \leq Q_1$$

The following simple procedure is proposed for carrying out the test:

1. Specify level of test, α (α = probability of Type I error)
2. Determine $(Q_2 - Q_1)_C = B_N - B_\Delta$
3. Calculate e_u according to Equation 83. [Note: the b_r's should be $(1 - 2\alpha) \times 100\%$ confidence limits, where α is the specified level of the test.]
4. If $(Q_2 - Q_1)_C + e_u$ is ≤ 0, reject H_0; otherwise, accept H_0. (Note: If H_0 is accepted, the plant is considered out of compliance.)

Notice that compliance requires that the emissions from the original plant be reduced by the amount of emissions from the new capacity plus an additional amount, e_u, which compensates for the uncertainty in the data.

The two kinds of incorrect decisions that can be made using the above procedure are:

Type I Error: Reject H_0 when it is, in fact, true; i.e., decide that the plant is in compliance when it is, in fact, in violation.

Type II Error: Accept H_0 when it is, in fact, false; i.e., decide the plant is in violation when it is, in fact, in compliance.

The probabilities of making the two types of errors in decision making are as follows:

1. Probability of Type I error is $\leq \alpha$.
2. Probability of Type II error is a function of α, $(Q_2 - Q_1)_C$, e_u, and e_ℓ. It must be determined individually for each specific problem.

Theoretically, the risk of a Type II error can be controlled by specifying the sample size to be used in obtaining emissions data. However, the effect of a Type II error is to require the plant to make an additional reduction in emissions. In practice, therefore, specification of the sample size involves an economic tradeoff between the cost of reducing uncertainty in the data and the potential cost of reducing emissions.

NUMERICAL EXAMPLE

In order to illustrate the test procedure, numerical values have been assigned to the variables involved in the test. These values are for illustration only, and do not correspond to actual field data.

1. Take $\alpha = 0.05$ [Note: the b_r's should be $(1 - 2\alpha) \times 100 = 90\%$ confidence limits.]

2. Let

$$\left.\begin{array}{l} b_{rN} = 6 \text{ g/s} \\ b_{r\Delta} = 8 \text{ g/s} \end{array}\right\} \qquad e_r = \sqrt{b^2_{rN} + b^2_{r\Delta}} = 10 \text{ g/s}$$

3. Let

$$\left.\begin{array}{l} b_{uN} = 4 \text{ g/s} \\ b_{\ell N} = 2 \text{ g/s} \\ b_{u\Delta} = 3 \text{ g/s} \\ b_{\ell\Delta} = 6 \text{ g/s} \end{array}\right\} \qquad \begin{array}{l} e_{su} = b_{uN} + b_{\ell\Delta} = 10 \text{ g/s} \\ e_{s\ell} = b_{\ell N} + b_{u\Delta} = 5 \text{ g/s} \end{array}$$

4. Then

$$e_u = e_r + e_{su} = 20 \text{ g/s}$$

$$e_\ell = e_r + e_{s\ell} = 15 \text{ g/s}$$

5. Let

$$\left.\begin{array}{l} B_N = 50 \text{ g/s} \\ B_\Delta = 65 \text{ g/s} \end{array}\right\} \qquad (Q_2 - Q_1)_C = B_N - B_\Delta = -15 \text{ g/s}$$

6. Uncertainty interval for $(Q_2 - Q_1)_T$ is then

$$\left(B_N - B_\Delta\right)_{-e_\ell}^{+e_u} = (-15)_{-15}^{+20} = (-30,5)$$

7. Since $(Q_2 - Q_1)_C + e_u = +5$ g/s, the null hypothesis is accepted; i.e., it is concluded that $(Q_2 - Q_1)_T > 0$ and the plant is out of compliance. Notice that if the uncertainty in the data were not taken into account, the opposite conclusion would be reached since the nominal value of $(Q_2 - Q_1)$ is -15 g/s.

8. An "effective" operating characteristic curve for the test is illustrated in Figure 9. This curve is obtained by treating the uncertainty interval (86) as an exact 90% confidence interval. This approximate treatment indicates a relatively large risk of a Type II error in this particular example ($P_{\text{Type II}} = 0.79$).

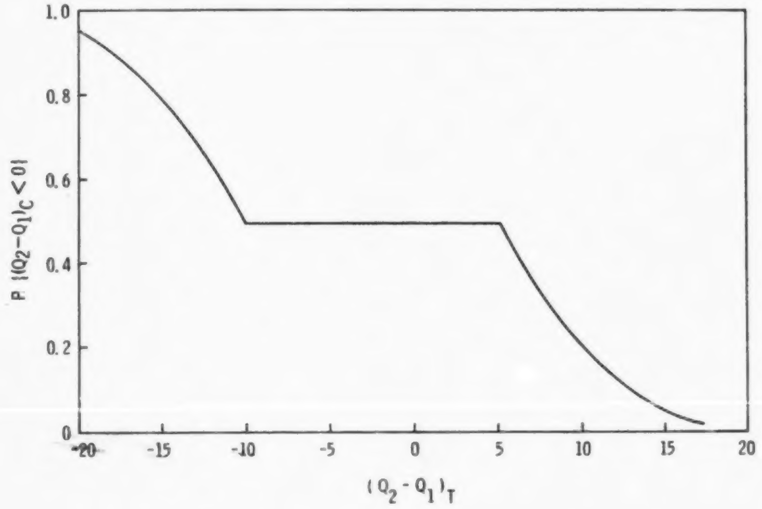


Figure 9. "Effective" operating characteristic curve for example problem.

SUMMARY

The data required to quantify the uncertainty in the emission rate difference before and after plant expansion consist of random and systematic error bounds on the following quantities:

- Emission rates from each emission point in the new unit.

- Emission rates before and after expansion from each emission point in the original unit which has different emissions after expansion.

The above data can be used to perform a simple, approximate statistical test to determine whether the plant is in compliance with offset policy. The essence of this procedure is the following: In order to ensure that air quality is not degraded, emissions from the original plant must be reduced by an amount sufficient to compensate not only for the emissions from the added capacity, but also for the uncertainty in emissions data.

SECTION 8

COMPARISON OF ALTERNATIVE CONTROLS: REASONABLY AVAILABLE CONTROL TECHNOLOGY

INTRODUCTION

The problem to be addressed in this section involves comparison of the control efficiencies of two alternative technologies. It is assumed that a plant currently employs a given control technique with a nominal (measured) control efficiency which is less than that of reasonably available control technology (RACT) for the given application. In order that the plant be required to adopt RACT, it should be established that RACT is indeed superior to the installed control technology. Since neither of the two efficiencies can be known with certainty, this determination should not be based on the nominal efficiencies alone.

The purpose of this section is to formulate a statistical test to serve as the basis for deciding whether RACT is superior to a given installed control technique. Only the simplest case, a single pollutant emitted from a single source, is considered.

GOVERNING EQUATIONS

The control efficiency, ε_1, of the installed device is given by

$$\varepsilon_1 = \frac{Q_1 - Q_2}{Q_1} = 1 - \frac{Q_2}{Q_1} \qquad (87)$$

where Q_1 and Q_2 are the uncontrolled and controlled emission rates, respectively. Letting ε_2 denote the control efficiency of RACT, we have

$$\varepsilon_2 - \varepsilon_1 = \varepsilon_2 + \frac{Q_2}{Q_1} - 1 \qquad (88)$$

For the purpose of this example, it is assumed that ε_1 is measured by measuring Q_1 and Q_2. The value of ε_2 is assumed to have been determined previously on a similar but different installation. Therefore, the measured value of ε_2, together with its associated random and systematic uncertainties, is assumed to be given.

We now wish to develop equations for the total uncertainty in the difference $(\varepsilon_2 - \varepsilon_1)$. Let B_1 and B_2 be the measured values of Q_1 and Q_2 with associated random uncertainties b_{r1} and b_{r2}. Let $(\varepsilon_2)_C$ be the given nominal value of ε_2 with associated random uncertainty β_r. Then the nominal value $(\varepsilon_2 - \varepsilon_1)_C$ of the control efficiency difference is given by

$$(\varepsilon_2 - \varepsilon_1)_C = (\varepsilon_2)_C + \frac{B_2}{B_1} - 1 \qquad (89)$$

The associated random uncertainty, e_r, is obtained using error propagation formulas for addition and division:

$$e_r = \sqrt{\beta_r^{\,2} + \left(\frac{B_2}{B_1}\right)^2 \left[\left(\hat{b}_{r1}\right)^2 + \left(\hat{b}_{r2}\right)^2\right]} \qquad (90)$$

where $\hat{b}_{r1} = b_{r1}/B_1$ and $\hat{b}_{r2} = b_{r2}/B_2$.

Considering next the systematic uncertainty, let

$$B_1{}^{+b_{u1}}_{-b_{\ell 1}},\ B_2{}^{+b_{u2}}_{-b_{\ell 2}},\ (\varepsilon_2)_C{}^{+\beta_u}_{-\beta_\ell},\ \text{and}\ (\varepsilon_2 - \varepsilon_1)_C{}^{+e_{su}}_{-e_{s\ell}}$$

be systematic error bounds for Q_1, Q_2, ε_2 and $(\varepsilon_2 - \varepsilon_1)$. Using addition and division formulas for propagation of systematic errors yields:

$$e_{su} = \beta_u + \left(\frac{B_2}{B_1}\right)\left(\frac{\hat{b}_{u2} + \hat{b}_{\ell 1}}{1 - \hat{b}_{\ell 1}}\right) \qquad (91)$$

$$e_{s\ell} = \beta_\ell + \left(\frac{B_2}{B_1}\right)\left(\frac{\hat{b}_{\ell 2} + \hat{b}_{u1}}{1 + b_{u1}}\right) \qquad (92)$$

where "^" again denotes relative uncertainty.

The total uncertainty in $(\varepsilon_2 - \varepsilon_1)$ is the sum of the random and systematic components. From Equations 79, 80, and 81 we obtain:

$$e_u = \sqrt{\beta_r^{\,2} + \left(\frac{B_2}{B_1}\right)^2 \left[\left(\hat{b}_{r1}\right)^2 + \left(\hat{b}_{r2}\right)^2\right]} + \beta_u + \left(\frac{B_2}{B_1}\right)\left(\frac{\hat{b}_{u2} + \hat{b}_{\ell 1}}{1 - \hat{b}_{\ell 1}}\right) \quad (93)$$

$$e_\ell = \sqrt{\beta_r^{\,2} + \left(\frac{B_2}{B_1}\right)^2 \left[\left(\hat{b}_{r1}\right)^2 + \left(\hat{b}_{r2}\right)^2\right]} + \beta_\ell + \left(\frac{B_2}{B_1}\right)\left(\frac{\hat{b}_{\ell 2} + \hat{b}_{u1}}{1 + \hat{b}_{u1}}\right) \quad (94)$$

The uncertainty interval for $(\varepsilon_2 - \varepsilon_1)$ is given by:

$$(\varepsilon_2 - \varepsilon_1)_C - e_\ell \le (\varepsilon_2 - \varepsilon_1)_T \le (\varepsilon_2 - \varepsilon_1)_C + e_u \quad (95)$$

where $(\varepsilon_2 - \varepsilon_1)_T$ represents the true value of the control efficiency difference, $(\varepsilon_2 - \varepsilon_1)_C$ is given by Equation 89, and e_u and e_ℓ are given by Equations 93 and 94.

TEST OF HYPOTHESIS

Inequality 95 forms the basis for a statistical test of hypothesis. The idea behind the test is the following: if the lower uncertainty limit $(\varepsilon_2 - \varepsilon_1)_C - e_\ell$ exceeds zero, then we can conclude that $(\varepsilon_2 - \varepsilon_1)_T$ is greater than zero, i.e., that RACT has a higher efficiency than the installed device. If the lower uncertainty limit is less than or equal to zero, then we cannot conclude that RACT is better.

The null hypothesis for this case is that RACT is no better than the installed control device, i.e.,

$$H_0: \quad (\varepsilon_2 - \varepsilon_1)_T \le 0$$

The alternative hypothesis is that RACT is better; i.e.,

$$H_A: \quad (\varepsilon_2 - \varepsilon_1)_T > 0$$

The test is formally carried out as follows:

1. Specify level of test, α (α = probability of Type I error).
2. Calculate $(\varepsilon_2 - \varepsilon_1)_C$ from Equation 89.
3. Calculate e_ℓ from Equation 94. [Note that the random uncertanties should be $(1 - 2\alpha) \times 100\%$ confidence limits.]
4. If $(\varepsilon_2 - \varepsilon_1)_C - e_\ell > 0$, reject H_0 (conclude RACT is better); otherwise accept H_0 (conclude RACT is not better).

The two types of incorrect decisions that can be made using the above test are:

Type I Error: Conclude RACT is better when in fact it is not better.

Type II Error: Conclude RACT is not better when, in fact, it is better.

NUMERICAL EXAMPLE

In order to illustrate the test procedure, numerical values have been assigned to the variables involved in the test. These values are for illustration only, and do not represent actual field data.

1. Take $\alpha = 0.05$ [Note: the random uncertainties should be $(1 - 2\alpha) \times 100 = 90\%$ confidence limits.]

2. Let $B_2/B_1 = 0.10$; i.e., the nominal efficiency of the installed device is $(\varepsilon_1)_C = 0.90$.

3. Let the nominal efficiency of RACT be $(\varepsilon_2)_C = 0.95$. Then $(\varepsilon_2 - \varepsilon_1)_C = 0.05$.

4. $\hat{b}_{r1} = \hat{b}_{r2} = 0.20$

   $\hat{b}_{u1} = \hat{b}_{\ell 1} = \hat{b}_{u2} = \hat{b}_{\ell 2} = 0.1$

   $\beta_r = 0.01$

   $\beta_u = \beta_\ell = 0$

Then,

$$e_\ell = \sqrt{\beta_r^{\,2} + \left(\frac{B_2}{B_1}\right)^2 \left[\left(\hat{b}_{r1}\right)^2 + \left(\hat{b}_{r2}\right)^2\right]} + \beta_\ell + \left(\frac{B_2}{B_1}\right)\left(\frac{\hat{b}_{\ell 2} + \hat{b}_{u1}}{1 + \hat{b}_{u1}}\right)$$

$$= 0.048$$

5. $(\varepsilon_2 - \varepsilon_1)_C - e_\ell = 0.05 - 0.048 = 0.002 > 0$

Therefore, we reject the null hypothesis and conclude that RACT has a higher control efficiency than the installed device.

SUMMARY

Before a plant is required to install RACT in preference to a previously installed alternative control method, the superiority of RACT should be demonstrated in light of uncertainties in emissions data. A simple approximate statistical test can be employed for this purpose. The essence of the test is the following: in order to demonstrate superiority, the RACT control efficiency is required to be sufficiently high that the difference in control efficiencies is statistically significant for the given level of uncertainty in the data.

The information required to perform the test consists of the random and systematic components of uncertainty in the two control efficiencies to be compared. Usually, this information will be derived from corresponding uncertainties in controlled and uncontrolled emission rates.

SECTION 9

CONCLUSIONS AND RECOMMENDATIONS

The principles of decision analysis can be used to form a sound basis for evaluating the uncertainties in environmental data as they relate to decision making. These principles have been applied to two numerical indices (source severity and national emissions burden) to develop guideposts for use as an aid in environmental decision making. However, these principles cannot be used in a rigorous fashion due to the complexity of the problems to be considered and due to the subjective nature of environmental decision making. Seldom, if ever, will an environmental decision be made only on the basis of a given numerical index (e.g., source severity). In general, other experimental data; social, political, and economic considerations; and human judgment all enter into the decisionmaking process. The index is only one piece of information which the decision maker may utilize in arriving at a decision. The index provides only technical information; value judgments must be made by the decision maker (in this case, IERL). This fact does not alter the basic approach used to analyze the uncertainty associated with the index. Therefore, in this report, it is convenient to think of the index alone as determining the outcome of the decision-making process. In practice, the index may not be used as the sole basis for decision making; IERL may consider other factors.

CONCLUSIONS

The following conclusions pertain to the Source Assessment Program:

- The stochastic decision approach is more than the deterministic approach, to uncertainties in emission factors, hence, it is the preferred approach.

- The critical value of the calculated source severity should be set at 0.05 or 0.08.

- The critical value for national emissions burden should be set at 0.05%.

- The following guidelines for planning sampling and analytical procedures should be observed:

Criteria pollutants--The random uncertainty in emission factor should be less than ±23%. If this restriction requires an impractically large number of samples, sufficient samples should be collected to maintain the random uncertainty below ±70%.

Noncriteria pollutants--The random uncertainty in emission factor should be less than ±23%. If this restriction requires an impractically large number of samples, the guideline for Cases B_1 and B_2 should be followed.

- In general, there will be a large risk of a Type II error (see Section 6) in the screening procedure no matter how large the value of the calculated source severity. This situation results from the dominant effect of systematic uncertainties associated with sampling and analytical procedures, dispersion modeling, and health effects information. This situation can be ameliorated only by reducing these uncertainties (e.g., by using more detailed modeling techniques or through a program of fundamental health effects research) or by devising a decision index other than source severity which will circumvent these uncertainties.

- Eliminating the uncertainty associated with dispersion modeling may be possible by appropriately defining the true (deterministic) source severity, S_T. Experimental data presented in Appendix I indicate that the simple Gaussian dispersion model correctly predicts the ensemble median ground level concentration under Class C stability conditions. Thus, by using this ensemble median value in the definition of S_T, the simple Gaussian equation could be used with correction factor (γ) equal to unity. In effect, the condition $\gamma = 1.0$ would be specified by the definition of S_T.

- Results of the source severity simulation presented in Appendix H indicate that the definition suggested in the preceeding conclusion for the deterministic source severity would be consistent with the stochastic approach to source severity. In particular, results indicate that the deterministic severity so defined would yield a worst case value in the sense that higher severities would be expected to occur only infrequently, on the order of once per year.

- Based on experimental data cited in Appendix I, the distribution of ground level concentration under Class C stability and fixed source conditions is approximately log-normally distributed, with 98% of the values falling within a factor of three of the median value. As noted above, the data indicate that the median concentration is correctly predicted by the simple Gaussian dispersion model.

RECOMMENDATIONS

The following recommendations are made for conducting the Source Assessment Program:

- The critical value of source severity for criteria pollutants should be set at 0.05.

- For noncriteria pollutants, the most consistent and logically defensible approach would be to estimate "acceptable" concentration using the mean TLV conversion factor for criteria pollutants (i.e., take F = TLV/21.4) and set the critical value of source severity at 0.01. The net effect would be to reduce the source severity (as originally specified for the Source Assessment Program) by a factor of about 15, and to reduce the original lower cut point by a factor of 10 from 0.1 to 0.01.

- For criteria pollutants, sampling and analytical strategies should be planned so as to maintain the random uncertainty in emission factor below ±23% if practicable; otherwise, below ±70%.

- For noncriteria pollutants, the minimum number (generally three) of samples required to obtain a valid estimate of the random uncertainty in emission factor should be collected and analyzed. This procedure is consistent with the second recommendation above, concerning interpretation and estimation of "acceptable" concentration.

The following recommendations are made for additional related work:

- The present study has served to underscore the need for additional information in the following areas:

  - Dispersion modeling--a simple method is needed to accurately predict the dispersion of airborne pollutants.

  - A vastly expanded data base is needed relating pollutant exposure to human health effects (extrapolation of animal data to humans).

  - Much more information is needed concerning the possible bias associated with individual sampling and analytical techniques used to measure pollutant emissions.

It is therefore recommended that additional research be undertaken in each of the above areas.

- The present study, including the above recommendations, should be considered as a first step toward establishing a coherent procedure for environmental decision making. It is recommended that a followup study be undertaken to investigate the applicability of other principles of decision theory, such as the principle of maximum expected utility, to the types of problems considered in this report. The present study is not the ultimate answer to analysis of the uncertainties in environmental data.

REFERENCES

1. Savage, L. J. The Foundations of Statistics. John Wiley & Sons, Inc., New York, New York, 1954.

2. Lindley, D. V. Making Decisions. John Wiley & Sons, Ltd., London, 1971.

3. Freund, J. E. Mathematical Statistics. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962. 390 pp.

4. TLVs® Threshold Limit Values for Chemical Substances and Physical Agents in the Workroom Environment with Intended Changes for 1976. American Conference of Governmental Industrial Hygienists, Cincinnati, Ohio, 1976. 94 pp.

5. Turner, D. B. Workbook of Atmospheric Dispersion Estimates. Public Health Service Publication No. 999-AP-26, U.S. Department of Health, Education, and Welfare, Cincinnati, Ohio, 1969. 84 pp.

6. Pasquill, F. Atmospheric Diffusion, Second Edition. John Wiley & Sons, Inc. (Halsted Press), New York, New York, 1974. 429 pp.

7. Cheremisinoff, P. N., and A. C. Morresi. Predicting Transport and Dispersion of Air Pollutants from Stacks. Pollution Engineering, 9(3):3, 26, 1977.

8. Hughes, T. W., R. B. Reznik, R. W. Serth, and Z. S. Khan. Source Assessment Methodology. Contract 68-02-1874, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, January 7, 1976. 33 pp.

9. Serth, R. W., and T. W. Hughes. Source Assessment: Carbon Black Manufacture. Contract 68-02-1874, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina. (Preliminary document submitted to the EPA by Monsanto Research Corporation, December 1975.) 145 pp.

10. Eimutis, E. C., B. J. Holmes, and L. B. Mote. Source Assessment: Severity of Stationary Air Pollution Sources--A Simulation Approach. EPA-600/2-76-032e, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, July 1976. 119 pp.

11. 1972 National Emissions Report; National Emissions Data System (NEDS) of the Aerometric and Emissions Reporting System (AEROS). EPA-450/2-74-012, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, June 1974. 422 pp.

12. Eisenhart, C. Expression of the Uncertainties of Final Results. Science, 160(3833):1201-1204, 1968.

13. Braddick, H. J. J. The Physics of Experimental Method. John Wiley & Sons, Inc., New York, New York, 1954. 404 pp.

14. Volk, W. Applied Statistics for Engineers. McGraw-Hill Book Company, New York, New York, 1958. 354 pp.

15. Beers, Y. Introduction to the Theory of Error. Second Edition. Addison-Wesley Publishing Co., Inc., Reading, Massachusetts, 1957.

16. Pearson, E. S., and H.O. Hartley. Biometrika Tables for Statisticians, Third Edition, Volume 1. Cambridge University Press, New York, New York, 1966. 264 pp.

17. Welch, B. L. The Generalization of "Student's" Problem When Several Different Population Variances Are Involved. Biometrika, 34:28-35, 1947.

18. Natrella, M. G. Experimental Statistics. National Bureau of Standards Handbook 91, U.S. Department of Commerce, Bureau of Standards, Washington, D.C., 1963. 504 pp.

19. Fisher, R. A., and F. Yates. Statistical Tables for Biological, Agricultural, and Medical Research. Oliver and Boyd, London, 1963. 356 pp.

20. Jenson, V. G., and G. V. Jeffreys. Mathematical Methods in Chemical Engineering. Academic Press, New York, New York, 1963. pp. 356-360.

21. Eimutis, E. C., and M. G. Konicek. Derivations of Continuous Functions for the Lateral and Vertical Atmospheric Dispersion Coefficients. Atmospheric Environment, 6(11):859-863, 1972.

22. Tadmor, J., and Y. Gur. Analytical Expressions for the Vertical and Lateral Dispersion Coefficients in Atmospheric Diffusion. Atmospheric Environment, 3:688-689, 1969.

23. Martin, D. O., and J. A. Tikvart. A General Atmospheric Diffusion Model for Estimating the Effects on Air Quality of One or More Sources. Presented at the 61st Annual Meeting of the Air Pollution Control Association, St. Paul, Minnesota, June 23-27, 1968. 18 pp.

24. Pasquill, F. Atmospheric Dispersion Parameters in Gaussian Plume Modelling, Part II, Possible Requirements for Change in the Turner Workbook Values. EPA-600/4-76-030b (PB 258 036), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, June 1976. 43 pp.

25. Code of Federal Regulations, Title 42--Public Health, Chapter IV--Environmental Portection Agency, Part 410--National Primary and Secondary Ambient Air Quality Standards, April 28, 1971. 16 pp.

26. Briggs, G. A. Plume Rise. AEC Critical Review Series. U.S. Atomic Energy Commission, Division of Technical Information Extension, Oak Ridge, Tennessee, 1969. 81 pp.

27. Moses, H., and M. R. Kraimer. Plume Rise Determination--A New Technique Without Equations. Journal of the Air Pollution Control Association, 22(8):621, 1972.

28. Thomas, F. W., S. B. Carpenter, and W. C. Colbaugh. Plume Rise Estimates for Electric Generating Stations. Journal of the Air Pollution Control Association, 20(3):170, 1970.

29. McGuire, T., and K. C. Knoll. Relationship Between Concentrations of Atmospheric Pollutants and Averaging Times. Atmospheric Environment, 5(5):291, 1971.

30. Handy, R., and A. Schindler. Estimation of Permissible Concentrations of Pollutants for Continuous Exposure. EPA-600/2-76-155 (PB 253 959), U.S. Environmental Protection Agency, Washington, D.C., June 1976. 136 pp.

31. Cornfield, J. Carcinogenic Risk Assessment. Science, 198 (4318):693-699, 1977.

32. Montgomery, T. L., and J. H. Coleman. Empirical Relationships Between Time-Averaged SO_2 Concentrations. Environmental Science and Technology, 9(10):953-957, 1975.

33. Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for Determination of Stack Gas Velocity and Volumetric Flow Rate in Conjunction with EPA Method 5. EPA-650/4-74-033 (PB 240 342), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 40 pp.

34. Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for Stack Gas Analysis and Determination of Moisture Fraction with Use of Method 5. EPA-650/4-73-026 (PB 236 929), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974.

35. Smith, F., and J. Buchanan. IERL-RTP Data Quality Manual. EPA-600/2-76-159, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1976.

36. Hamil, H. F., and D. E. Camann. Collaborative Study of Method for the Determination of Particulate Matter Emissions from Stationary Sources (Portland Cement Plants). EPA-650/4-74-029 (PB 237 346), U. S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 54 pp.

37. Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for the Determination of Particulate Matter Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators). EPA-650/4-74-021 (PB 234 150), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 36 pp.

38. Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for the Determination of Particulate Matter Emissions from Stationary Srouces (Municipal Incinerators). EPA-650/4-74-022 (PB 234 151), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 37 pp.

39. Hamil, H. F., and D. E. Camann. Collaborative Study of Method for the Determination of Sulfur Dioxide Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators). EPA-650/4-74-024 (PB 238 293), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1973. 64 pp.

40. Hamil, H. F., and D. E. Camann. Collaborative Study of Method for the Determination of Nitrogen Oxide Emissions from Stationary sources (Fossil Fuel-Fired Steam Generators). EPA-650/4-74-025 (PB 238 555), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1973. 102 pp.

41. Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for the Determination of Nitrogen Oxide Emissions from Stationary Sources (Nitric Acid Plants). EPA-650/4-74-028 (PB 236 930), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 41 pp.

42. Constant, P. C., G. Scheil, and M. C. Sharp. Collaborative Study of Method 10--Reference Method for Determination of Carbon Monoxide Emissions from Stationary Sources--Report of Testing. EPA-650/4-75-001 (PB 241 284), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1975.

43. Constant, P. C., and M. C. Sharp. Collaborative Study of Method 104--Reference Method for Determination of Beryllium Emissions from Stationary Sources. EPA-650/4-74-023 (PB 245 011), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 94 pp.

44. Title 40--Protection of Environment. Chapter 1--Environmental Protection Agency--Part 60--Standards of Performance for New Stationary Sources. Subchapter C--Air Programs. Method 10--Determination of Carbon Monoxide Emissions from Stationary Sources. Federal Register, 39(47):9319-9321, 1974.

45. Guzewich, D. C., and W. J. B. Pringle. Validation of the EPA-PTMTP Short-Term Gaussian Dispersion Model. Journal of the Air Pollution Control Association, 27(6):540-542, 1977.

46. Scriven, R. A. Variability and Upper Bounds for Maximum Ground Level Concentrations. Philosophical Transactions of the Royal Society of London, Series A, 265:209-220, 1969.

47. Weber, A. H. Atmospheric Dispersion Parameters in Gaussian Plume Modelling, Part I, Review of Current Systems and Possible Future Developments. EPA-600/4-76-030a (PB 257 893), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, July 1976. 58 pp.

48. Shiruaikar, V. V., and P. R. Patel. Long Term Statistics of Peak/Mean Concentrations From A Point Sources. Atmospheric Environment, 11(4):387-389, 1977.

49. Serth, R. W., and T. W. Hughes. Source Assessment: Carbon Black Manufacture. EPA-600/2-77-107K, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, October 1977. 244 pp.

50. Standard for Metric Practice. ANSI/ASTM Designation E 380-76$^{\epsilon}$, IEEE Std 268-1976, American Society for Testing and Materials, Philadelphia, Pennsylvania, February 1976. 37 pp.

APPENDIX A

ACCURACY, ERROR, AND UNCERTAINTY

In the analysis of errors in experimental data, it is important to distinguish the following concepts: 1) accuracy of the measurement process, 2) error in the measured value, and 3) uncertainty in our knowledge of the true value which is measured. The relationship between these concepts is indicated schematically in Figure A-1.

MEASUREMENT PROCESS		MEASURED VALUE		KNOWLEDGE OF TRUE VALUE
ACCURACY	→	(TOTAL) ERROR	→	(TOTAL) UNCERTAINTY
PRECISION	→	RANDOM ERROR	→	RANDOM UNCERTAINTY
BIAS	→	SYSTEMATIC ERROR	→	SYSTEMATIC UNCERTAINTY

Figure A-1. Schematic representative of relationship between accuracy, error, and uncertainty and between the components of each term.

The terms accuracy, precision, and bias refer to the measurement process itself (12). Accuracy refers to the closeness between true and measured values which is characteristic of a given measurement process. Accuracy (strictly speaking, inaccuracy) is composed of two parts: precision (strictly speaking, imprecision) and bias. Precision is a measure of the closeness together, or lack of scatter, in successive independent measurements when all controllable variables are held fixed. Scatter in the measurements is due to fluctuations in values of variables which are not controlled in the measurement process. Bias is the magnitude and direction of the tendency of the measurement process to measure something other than what is intended. Failure to maintain isokinetic conditions during particulate sampling is a typical example of bias in a measurement process.

The inaccuracy of the measurement process results in errors in the measured values. The total error in a measurement is the magnitude and sign of the deviation of the measured value from the true value.

---

(12) Eisenhart, C. Expression of the Uncertainties of Final Results. Science, 160(3833):1201-1204, 1968.

The total error can be decomposed into a random component and a systematic component. The random error in the measured value results from the imprecision of the measurement process, while the systematic error results from the bias in the measurement process. The decomposition of total error into random and systematic components is illustrated in Figure A-2. When the mean of the population of all possible measured values coincides with the true value in Figure A-2, the measurement process is said to be unbiased.

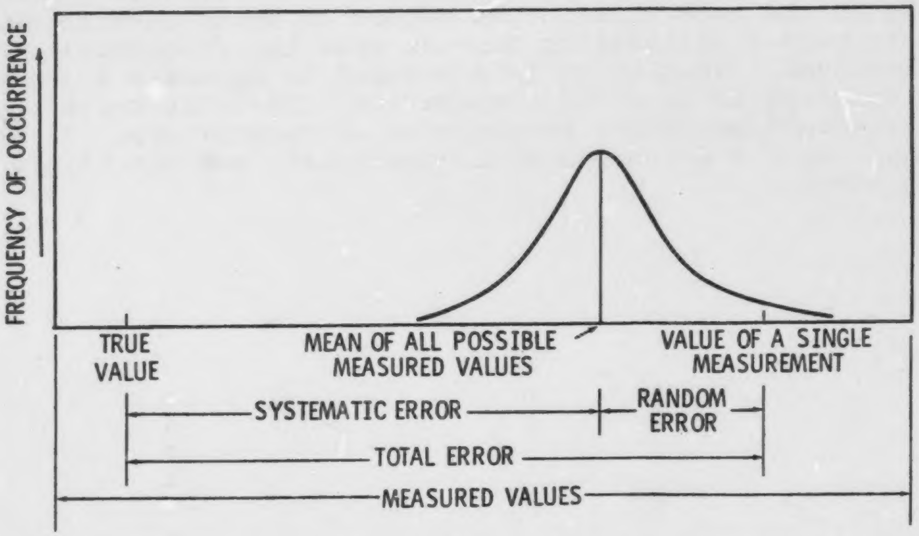


Figure A-2. Decomposition of total error into random and systematic components.

In practice, the magnitude of the error in the measured value is not known. If it were known, the measured value could simply be corrected by the amount of the error to obtain the true value. The best that can usually be done is to place reasonable bounds on the possible error. In the case of random errors, the bounds can be estimated statistically, and the results expressed in the form of a confidence interval about the measured value. In the case of systematic errors, bounds must be estimated based on a knowledge of the measurement process. For example, in field sampling work, the measurement method can be tested on standard samples to determine the bias under varying conditions. The largest value of the bias so obtained may then be used to estimate a bound on the systematic error under field conditions. Again, these results can be expressed in the form of an interval about the measured value.

Thus, the estimation of bounds for the error in a measured value results in an interval about the measured value within which the true value can reasonably be expected to lie; i.e., an interval of uncertainty. Uncertainty, then, refers to our lack of knowledge of the true value of a quantity due to imperfect data;

i.e., due to error in the measured value. The total uncertainty can be decomposed into a random component, which is due to the random error in the measured value, and a systematic component, which is due to the systematic error in the measured value. The uncertainty is quantified by estimating bounds on the random and systematic errors.

When measured values are used in a calculation, the corresponding uncertainties result in an uncertainty of the calculated quantity in relation to the true value. The manner in which uncertainties propagate through a calculation depends upon the functional relationship involved. This topic is discussed in Appendix B and is generally referred to as error propagation. Strictly speaking, however, it should be called propagation of uncertainty. In this report, then, we are concerned with uncertainty and its effect on decision making.

APPENDIX B

USE AND INTERPRETATION OF ERROR PROPAGATION FORMULAS

Any experimentally determined quantity has associated with it some degree of error for which bounds must be determined in order for the measurement to be meaningful. Indeed, a measured value for which no error bounds can be determined is essentially useless. When the measured value is subsequently used in a calculation, its associated error results in an error in the computed value. One is then faced with the problem of determining error bounds on the computed value in terms of errors associated with input values.

The manner in which errors (strictly speaking, uncertainties[a]) propagate in a calculation depends upon the functional relationship between variables and upon the type (random or systematic[a]) of error involved. In this appendix, the basic formulas for propagating the two types of errors are reviewed, and a number of example problems are presented which illustrate their use and limitations. Some brief derivations are included to point out underlying assumptions and to indicate the extension to arbitrary functional relationships.

As pointed out in Appendix A, the subject of the present section should more properly be termed "propagation of uncertainty" rather than "propagation of error." However, in order to adhere to convention as well as for conciseness of expression, the term "error" is used in this section to denote both exact error (which is useful in theoretical discussions) and uncertainty (which is useed in practice). The intended meaning will usually be clear from the context. When it is not clear, the meaning will be indicated parenthetically.

RANDOM ERRORS

Error propagation formulas for random errors are listed in Table B-1 for the four basic arithmetic operations. These formulas are special cases of the general formula also given in Table B-1. The formulas give the confidence interval for the appropriate function of two independent variables in terms of the confidence intervals for the two independent variables. The extension to cases involving more than two independent variables should be evident from the form of the general equation listed in the table. These formulas as well as others are given in various forms in the literature (see, for example, References 13, 14, and 15). All of the formulas are based on the following equation for the variance of a linear function of n statistically independent random variables (3):

---

[a] See discussion in Appendix A.

$$\sigma^2_y = (C_1)^2 \sigma^2 x_1 + (C_2)^2 \sigma^2 x_2 + \ldots + (C_n)^2 \sigma^2 x_n \qquad \text{(B-1)}$$

where $y = C_1x_1 + C_2x_2 + \ldots + C_nx_n$, and the C_i are constants.

If the x_i are not statistically independent, then the equation for σ^2_y contains the covariances of the x variables and hence the equations in Table B-1 are not valid. A derivation of the error propagation formula for the sum of two variables is first given, and then its extension to nonlinear relationships is indicated.

TABLE B-1. ERROR PROPAGATION FORMULAS FOR RANDOM ERRORS

Operation	Error propagation formula
Addition, $x_1 + x_2$	$A + B \pm \sqrt{a^2 + b^2}$
Subtraction, $x_1 - x_2$	$A - B \pm \sqrt{a^2 + b^2}$
Multiplication, x_1x_2	$AB \pm \sqrt{B^2a^2 + A^2b^2}$
Division, x_1/x_2	$(A/B) \pm \sqrt{\frac{a^2}{B^2} + \frac{A^2}{B^4} b^2}$
General case: $f(x_1, x_2)$	$f(A,B) \pm \sqrt{\left[\frac{\partial f(A,B)}{\partial x_1}\right]^2 a^2 + \left[\frac{\partial f(A,B)}{\partial x_2}\right]^2 b^2}$

NOTE.—A ± a and B ± b are confidence intervals for $\bar{x}_1$ and $\bar{x}_2$. The formulas give confidence intervals for the various mathematical operations performed with $\bar{x}_1$ and $\bar{x}_2$. The formulas are valid only when A and B are statistically independent.

---

(13) Braddick, H. J. J. The Physics of Experimental Method. John Wiley & Sons, Inc., New York, New York, 1954. 404 pp.

(14) Volk, W. Applied Statistics for Engineers. McGraw-Hill Book Company, New York, New York, 1958. 354 pp.

(15) Beers, Y. Introduction to the Theory of Error. Second Edition. Addison-Wesley Publishing Co., Inc., Reading, Massachusetts, 1957.

Derivation of Addition Formula

Let $y = x_1 + x_2$, where x_1 and x_2 are independent variables whose measurements are subject to random errors. The (exact) errors are assumed to be normally distributed with zero means and variances $\sigma^2_{x_1}$ and $\sigma^2_{x_2}$. Let A and B denote the estimated means of x_1 and x_2 determined from n_A measurements of x_1 and n_B measurements of x_2. Then $(1 - \alpha) \times 100\%$ confidence intervals for the true average values $\overline{x}_1$ and $\overline{x}_2$ are given by

$$\overline{x}_1 = A \pm Z_{1-\alpha/2}\,\sigma_A = A \pm Z_{1-\alpha/2}\frac{\sigma_{x_1}}{\sqrt{n_A}} = A \pm a \qquad \text{(B-2)}$$

$$\overline{x}_2 = B \pm Z_{1-\alpha/2}\,\sigma_B = B \pm Z_{1-\alpha/2}\frac{\sigma_{x_2}}{\sqrt{n_B}} = B \pm b \qquad \text{(B-3)}$$

where $Z_{1-\alpha/2}$ denotes the $(1 - \alpha/2)$ percentage point of the standard normal distribution, where $\sigma^2_A = \sigma^2_{x_1}/n_A$ and $\sigma^2_B = \sigma^2_{x_2}/n_B$ are the variances of the two sample means A and B, and where

$$a \equiv Z_{1-\alpha/2}\frac{\sigma_{x_1}}{\sqrt{n_A}} \qquad \text{(B-4)}$$

$$b \equiv Z_{1-\alpha/2}\frac{\sigma_{x_2}}{\sqrt{n_B}} \qquad \text{(B-5)}$$

A $(1 - \alpha) \times 100\%$ confidence interval for $\overline{y}$ is required. Since $\overline{y} = \overline{x}_1 + \overline{x}_2$, $\overline{y}$ is normally distributed (3), and the sample estimate of $\overline{y}$ is $C = A + B$. Using Equation B-1,

$$\sigma^2_C = \sigma^2_A + \sigma^2_B = \frac{\sigma^2_{x_1}}{n_A} + \frac{\sigma^2_{x_2}}{n_B} \qquad \text{(B-6)}$$

Thus, the $(1 - \alpha) \times 100\%$ confidence interval for $\overline{y}$ is

$$\overline{y} = C \pm Z_{1-\alpha/2}\,\sigma_C = A + B \pm Z_{1-\alpha/2}\sqrt{\frac{\sigma^2_{x_1}}{n_A} + \frac{\sigma^2_{x_2}}{n_B}}$$

$$= A + B \pm \sqrt{a^2 + b^2} \qquad \text{(B-7)}$$

When the variances of $\sigma^2_{x_1}$ and $\sigma^2_{x_2}$ are not known, they must be estimated by the sample variances

$$s^2_{x_1} = \frac{1}{n_A - 1} \sum_{i=1}^{n_A} \left(x_{1_i} - A\right)^2; \; s^2_{x_2} = \frac{1}{n_B - 1} \sum_{i=1}^{n_B} \left(x_{2_i} - B\right)^2 \quad \text{(B-8)}$$

The confidence intervals for $\bar{x}_1$ and $\bar{x}_2$ are then

$$A \pm t_{1-\alpha/2,\, n_A - 1} \frac{s_{x_1}}{\sqrt{n_A}}$$

and

$$B \pm t_{1-\alpha/2,\, n_B - 1} \frac{s_{x_2}}{\sqrt{n_B}}$$

where $t_{1-\alpha/2,\nu}$ denotes the $(1 - \alpha/2)$ percentage point of the Student t distribution with ν degrees of freedom. In this case, the above derivation is no longer valid because although

$$s^2_C = \frac{s^2_{x_1}}{n_A} + \frac{s^2_{x_2}}{n_B} \quad \text{(B-9)}$$

is an unbiased estimate of σ^2_C, it is not a sum-of-square estimate. Therefore, the random variable $(C - \bar{y})/s_C$ will not follow the t distribution. (The special case when $\sigma^2_{x_1} = \sigma^2_{x_2}$ and $n_A = n_B$ is an exception. See Example 3 below.) Thus in this case, the error propagation formulas are approximations which are accurate for large sample sizes n_A and n_B, for which the estimates $s^2_{x_1}$ and $s^2_{x_2}$ approach the true variances $\sigma^2_{x_1}$ and $\sigma^2_{x_2}$. However, it might be expected that the formulas will not be greatly in error except for sample sizes smaller than about 5, for which the percentage points of the t distribution deviate markedly from those of the normal distribution. This point is examined further in the examples given below.

Nonlinear Relations

When $y = f(x_1,x_2)$ and f is nonlinear, the above derivation does not apply because Equation B-1 no longer holds. However, when errors in measurements are small (or more precisely, when standard deviations σ_{x_1} and σ_{x_2} are small compared with mean values $\bar{x}_1$ and $\bar{x}_2$), the problem can be linearized. Thus, neglecting second-order terms,

$$C \equiv f(A,B) \simeq f(\bar{x}_1,\bar{x}_2) + \frac{\partial f(\bar{x}_1,\bar{x}_2)}{\partial x_1}(A - \bar{x}_1) + \frac{\partial f(\bar{x}_1,\bar{x}_2)}{\partial x_2}(B - \bar{x}_2) \quad \text{(B-10)}$$

Since this expression is linear in A and B and since to this degree of approximation the partial derivatives evaluated at $(\bar{x}_1, \bar{x}_2)$ and (A,B) are equal, Equation B-1 yields

$$\sigma^2_C = \left[\frac{\partial f(A,B)}{\partial x_1}\right]^2 \sigma^2_A + \left[\frac{\partial f(A,B)}{\partial x_2}\right]^2 \sigma^2_B \tag{B-11}$$

For example, when $y = f(x_1, x_2) = x_1 x_2$, then $\partial f(A,B)/\partial x_1 = B$ and $\partial f(A,B)/\partial x_2 = A$, so that

$$\sigma^2_C = B^2\sigma^2_A + A^2\sigma^2_B \tag{B-12}$$

The confidence interval for $\bar{y}$ is then

$$\bar{y} = C \pm Z_{1-\alpha/2}\, \sigma_C = AB \pm Z_{1-\alpha/2} \sqrt{B^2\sigma^2_A + A^2\sigma^2_B}$$

$$= AB \pm \sqrt{B^2a^2 + A^2b^2} \tag{B-13}$$

which is the multiplication formula given in Table B-1. In this manner, error propagation formulas can be derived for any functional relationship. However, the formulas will be accurate for only small errors.

Non-Normally Distributed Errors

When error distributions deviate significantly from normality, Chebyshev's theorem (3) can be used to estimate confidence intervals. This theorem states that for a random variable, x, which has a mean, μ, and standard deviation, σ, the probability that $|x - \mu| > k\sigma$ is less than $1/k^2$; i.e.,

$$P(|x - \mu| > k\sigma) < \frac{1}{k^2} \tag{B-14}$$

Thus, a $(1 - \alpha) \times 100\%$ confidence interval is given by

$$\mu \pm k\sigma$$

where $k\sigma = \frac{1}{\sqrt{\alpha}}$

Actually, the confidence level will be greater than $(1 - \alpha) \times 100\%$ since the above probability is strictly less than $1/k^2$. The important point is that this confidence interval is independent of the distribution of the random variable, x.

If Chebyshev's theorem is used to compute all confidence intervals, the derivations of error propagation formulas given above remain valid if k_α is substituted for $z_{1-\alpha/2}$. The formulas will be only approximately correct if sample variances are used to estimate error variances or if normal statistics are used to compute one confidence interval and Chebyshev's theorem is used to compute the other.

For a 95% confidence interval,

$$1 - \alpha = 0.95$$

$$\alpha = 0.05$$

$$k_\alpha = 4.47$$

This value of k_α is more than twice the corresponding value of $z_{1-\alpha/2} = 1.96$ for a normally distributed variate. For this reason, Chebyshev's theorem is seldom used to compute confidence intervals. It is recommended that it be used only when error distributions are known to be highly non-normal.

Examples: Random Errors

Example 1--

The following data on atmospheric emission factors for ammonia emitted from ammoniation and granulation plants were obtained from various published sources.

TABLE B-2. AMMONIA EMISSIONS FROM AMMONIUM NITRATE PLANTS

	Ammoniator-granulator	Dryer and cooler
Number of measurements, n	7	16
Mean emission factor, g/kg	0.503	0.316
s, g/kg	0.564	0.262
$s/\sqrt{n}$	0.213	0.065
$t_{0.975}$, n - 1	2.447	2.131
95% Confidence interval for mean	±0.521 (±104%)	±0.139 (±44%)

The emission factor for the entire plant is the sum of the values for the ammoniator-granulator and the dryer and cooler. We wish to compute the confidence interval for the emission factor for the entire plant and to compare the result obtained using the error propagation formula with that obtained by rigorous statistical methods.

- Using the error propagation formula,

$$A = 0.503,\ a = 0.521,\ B = 0.316,\ b = 0.139$$

Hence,

$$C = A + B = 0.819$$

$$c = \sqrt{a^2 + b^2} = \sqrt{0.271 + 0.019} = 0.539$$

The 95% confidence interval is thus 0.819 ± 0.539, or 0.819 ± 66%.

- The confidence interval can be approximated using the method of Welch (16, 17). Actually, an approximation to Welch's method is used which is valid when sample sizes are not too small (18). The following quantities are computed:

$$V_A \equiv \frac{s^2{}_{x_1}}{n_A} = 0.213^2 = 0.0454$$

$$V_B \equiv \frac{s^2{}_{x_2}}{n_B} = 0.065^2 = 0.0042$$

$$\sqrt{V_A + V_B} = 0.2228 = \text{"effective" standard deviation of mean}$$

$$f = \left[\frac{\left(V_A + V_B\right)^2}{\dfrac{V_A{}^2}{n_A + 1} + \dfrac{V_B{}^2}{n_B + 1}}\right] - 2 = \text{"effective" degrees of freedom}$$

$$f = 8.37 - 2 = 6.37$$

The confidence interval is computed using the t distribution with f' degrees of freedom, f' being the nearest integer to f, or f' = 6 in this case. Thus,

$$c = t_{0.975,6}\sqrt{V_A + V_B} = 2.447(0.2228) = 0.545$$

---

(16) Pearson, E. S., and H. O. Hartley. Biometrika Tables for Statisticians, Third Edition, Volume 1. Cambridge University Press, New York, New York, 1966. 264 pp.

(17) Welch, B. L. The Generalization of "Student's" Problem When Several Different Population Variances Are Involved. Biometrika, 34:28-35, 1947.

(18) Natrella, M. G. Experimental Statistics. National Bureau of Standards Handbook 91, U.S. Department of Commerce, Bureau of Standards, Washington, D.C., 1963. 504 pp.

The 95% confidence interval is then 0.819 ± 0.545 or 0.819 ± 67%. The agreement with the previous calculation is excellent.

- Suppose that it is known *a priori* that the variances $\sigma^2 x_1$ and $\sigma^2 x_2$ of the two error distributions are equal. (Actually, using the F-test, the difference in the two sample variances is not significant at the 0.05 level, but is significant at the 0.10 level.) In this case, the confidence interval can be computed rigorously using the t distribution. The method is essentially the same as that employed for comparing two means (21). The pooled estimate of the unknown variance is computed as follows:

  $$s^2{}_P = \frac{\left(n_A - 1\right)s^2{}_{x_1} + \left(n_B - 1\right)s^2{}_{x_2}}{n_A + n_B - 2} = 0.1399$$

  The estimated standard deviation of the mean is then

  $$s_P\sqrt{1/n_A + 1/n_B} = 0.3741\sqrt{1/7 + 1/16} = 0.1695$$

  The confidence limit is given by

  $$\begin{aligned} t_{0.975},\ n_A + n_B - 2\ s_P\sqrt{1/n_A + 1/n_B} &= t_{0.975,21}(0.1695) \\ &= 2.080(0.1695) \\ &= 0.353 \end{aligned}$$

  Hence, the confidence interval is 0.819 ± 0.353 or 0.819 ± 43%. It is seen that the additional information about the error variances can be used to obtain a smaller confidence interval.

Example 2--
The data are similar to those of Example 1, but the emission factors are for particulate matter rather than ammonia.

TABLE B-3. PARTICULATE EMISSIONS FROM AMMONIUM NITRATE PLANTS

	Ammoniator-granulator	Dryer and cooler
Number of measurement, n	2	12
Mean emission factor, g/kg	0.175	0.230
s, g/kg	0.070	0.173
$s/\sqrt{n}$	0.049	0.050
$t_{0.975}$, n - 1	12.706	2.201
95% Confidence limit for mean	0.623(356%)	0.110(48%)

This example represents an extreme case since only two measurements are available for the ammoniator-granulator.

- Using the error propagation formula,

$$A = 0.175,\ a = 0.623,\ B = 0.230,\ b = 0.110$$

Hence,

$$C = A + B = 0.405$$

$$c = \sqrt{a^2 + b^2} = 0.632$$

The 95% confidence interval is thus 0.405 ± 0.632 or 0.405 ± 156%.

- The approximation to Welch's method used in Example 1 is not valid in this case due to the small sample size involved ($n_A = 2$). On the other hand, the exact version of the method cannot be used because the percentage points of the appropriate distribution function have not been tabulated for degrees of freedom less than 8 (19). The approximate calculation is included here for completeness.

$$V_A = \frac{s^2_{x_1}}{n_A} = 0.00245$$

$$V_B = \frac{s^2_{x_2}}{n_B} = 0.00250$$

$$\sqrt{V_A + V_B} = 0.0704$$

$$f = \left[\frac{(V_A + V_B)^2}{\frac{V_A{}^2}{n_A + 1} + \frac{V_B{}^2}{n_B + 1}}\right] - 2 = 9.88 - 2 = 7.88$$

Therefore,

$$f' = 8$$

$$t_{0.975,8}\sqrt{V_A + V_B} = 2.306(0.0704) = 0.162$$

Hence, the confidence interval is 0.405 ± 0.162 or 0.405 ± 40%. The agreement here is quite poor, but no conclusion can be made since the accuracy of the present calculation is unknown.

- It is assumed that $\sigma^2_{x_1} = \sigma^2_{x_2}$. (Using the F test, the difference in the sample variances is not significant even at the 0.25 significance level.) Then

$$s^2_P = \frac{\left(n_A - 1\right)s^2_{x_1} + \left(n_B - 1\right)s^2_{x_2}}{n_A + n_B - 2} = 0.0278$$

$$s_P\sqrt{1/n_A + 1/n_B} = 0.0162$$

$$t_{0.975,\ n_A + n_B - 2}\ s_P\sqrt{1/n_A + 1/n_B} = 2.179(0.0162)$$

$$= 0.277$$

Hence, the confidence interval is 0.405 ± 0.277 or 0.405 ± 68%. Since the difference in the sample variances is not significant at the 0.25 level, the present calculation is probably the most accurate of the three methods in this case. Thus, it appears that error propagation formulas can yield very conservative error estimates when small sample sizes are involved.

Example 3--
It is desirable to investigate further the behavior of the formulas for small sample sizes since repeating an experiment more than once or twice is often impractical. As previously noted, when $y = x_1 + x_2$, the random variable

$$\frac{C - \overline{y}}{s_C}$$

follows the t distribution with 2(n - 1) degrees of freedom in the special case when $\sigma^2_{x_1} = \sigma^2_{x_2}$ and $n_A = n_B \equiv n$. Hence, the confidence limits for $\overline{y}$ are given by

$$\pm\ t_{1-\alpha/2,\ 2(n-1)}\ s_C$$

On the other hand, the addition formula yields the confidence limits

$$\pm\ t_{1-\alpha/2,\ n-1}\ s_C$$

The percent difference is thus

$$\frac{t_{n-1}\, s_C - t_{2(n-1)}\, s_C}{t_{2(n-1)}\, s_C}(100) = \frac{t_{n-1} - t_{2(n-1)}}{t_{2(n-1)}}(100)$$

where the confidence level subscripts have been dropped for convenience. This factor is tabulated below for the 95% and 99% (two-sided) confidence levels ($1 - \alpha/2 = 0.975$ and 0.995).

TABLE B-4. PERCENT DIFFERENCES IN SIZES OF APPROXIMATE AND EXACT CONFIDENCE INTERVALS

	$\frac{t_{n-1} - t_{2(n-1)}}{t_{2(n-1)}}(100)$	
n	95% Level	99% Level
2	191	541
3	55	116
4	30	58
5	25	37
10	8	13

It is clear from these values that error propagation formulas can yield very conservative results for smaller sample sizes. However, at the 95% level--the one most often used in practice--the difference does not exceed 25% for n = 5. A conservative error of this magnitude in the confidence limits should be acceptable for most purposes. In effect, the calculated confidence interval simply corresponds to a slightly higher significance level than the nominal (95%) level.

When $\sigma^2 x_1 \neq \sigma^2 x_2$, the accuracy of the error propagation formulas can be estimated by means of the Behrens-Fisher confidence interval, $d_{n-1}\, s_C$ (19). Aside from philosophical considerations, the Behrens-Fisher method may be disputed on the grounds that the actual confidence level is not exactly equal to the nominal level, but varies with the ratio σ_{x1}/σ_{x2} (16). However, the method should give a good indication of the accuracy of the error formulas. The percent error in the addition formula is given by

$$\frac{t_{n-1} - d_{n-1}}{d_{n-1}}(100)$$

where d_{n-1} is tabulated in Table VI of Reference 19. Since d_{n-1} depends on $s_A/s_B \equiv \tan\theta$, the extreme value of d_{n-1} for

---

(19) Fisher, R. A., and F. Yates. Statistical Tables for Biological, Agricultural, and Medical Research. Oliver and Boyd, London, 1963. 356 pp.

$0° \leq \theta \leq 90°$ is used to estimate the maximum error. The values are listed below for the 95% and 99% levels.

TABLE B-5. MAXIMUM PERCENT DIFFERENCES IN SIZES OF APPROXIMATE AND EXACT CONFIDENCE INTERVALS WHEN ERROR VARIANCES ARE UNEQUAL

	$\frac{t_{n-1} - d_{n-1}}{d_{n-1}}(100)$	
n	95% Level	99% Level
2	-29	-29
4	-1.8	+4.3
6	+0.35	+5.8

Significant errors occur for only the smallest sample sizes; however, the errors need not be conservative.

Example 4--
Derive error propagation formulas for the exponential and logarithmic operations. For the exponential operation,

$$y = f(x) = x^n \qquad \text{(B-15)}$$

$$f'(x) = nx^{n-1} \qquad \text{(B-16)}$$

$$C = f(A) = A^n \qquad \text{(B-17)}$$

and

$$\bar{x} = A \pm a \qquad \text{(B-18)}$$

where $a = z_{1-\alpha/2}\, \sigma_A$

Following the linearization procedure described above under "Nonlinear Relations,"

$$\sigma_C{}^2 = [f'(A)]^2(\sigma_A{}^2) \qquad \text{(B-19)}$$

$$\sigma_C = f'(A)\sigma_A \qquad \text{(B-20)}$$

$$\sigma_C = nA^{n-1}(\sigma_A) \qquad \text{(B-21)}$$

$$\bar{y} = C \pm z_{1-\alpha/2}(\sigma_C) \qquad \text{(B-22)}$$

$$\bar{y} = A^n \pm z_{1-\alpha/2}\, nA^{n-1}\left(\sigma_A\right) \quad \text{(B-23)}$$

$$\bar{y} = A^n \pm nA^{n-1}(a) \quad \text{(B-24)}$$

Symbolically, this relationship may be written as

$$(A \pm a)^n = A^n \pm nA^{n-1}(a) \quad \text{(B-25)}$$

For the logarithmic operation,

$$y = f(x) = \ln(x) \quad \text{(B-26)}$$

$$f'(x) = 1/x \quad \text{(B-27)}$$

$$C = f(A) = \ln(A) \quad \text{(B-28)}$$

Linearizing as before,

$$\sigma_C{}^2 = [f'(A)]^2\left(\sigma_A{}^2\right) \quad \text{(B-29)}$$

$$\sigma_C = f'(A)\sigma_A = \frac{1}{A}\left(\sigma_A\right) \quad \text{(B-30)}$$

Therefore,

$$\bar{y} = C \pm z_{1-\alpha/2}\left(\sigma_C\right) \quad \text{(B-31)}$$

$$= \ln(A) \pm z_{1-\alpha/2}\,\frac{1}{A}\left(\sigma_A\right) \quad \text{(B-32)}$$

$$\bar{y} = \ln(A) \pm \frac{a}{A} \quad \text{(B-33)}$$

Symbolically, this result may be written

$$\ln(A \pm a) = \ln\left(A\right) \pm \frac{a}{A} \quad \text{(B-34)}$$

The above formulas may also be obtained as special cases of the general error propagation formula given in Table B-1.

SYSTEMATIC ERRORS

Error propagation formulas for systematic errors are given in Table B-6 for the four basic arithmetic operations. The first two formulas are valid in general, while the latter two are restricted by the conditions $a < |A|$ and $b < |B|$. In all cases, the two variables as well as the corresponding errors are assumed to be functionally independent. The first two formulas together with the linearized versions of the multiplication and division formulas have been reported in the literature (20). A derivation of the multiplication formula is given here to demonstrate the method.

TABLE B-6. ERROR PROPAGATION FORMULAS FOR SYSTEMATIC ERRORS

Operation	Lower bound	Upper bound
Addition	$A + B - (a + b)$	$A + B + (a + b)$
Subtraction	$A - B - (a + b)$	$A - B + (a + b)$
Multiplication	$AB + \text{sgn}(AB)ab - (a\|B\| + b\|A\|)$	$AB + \text{sgn}(AB)ab + (a\|B\| + b\|A\|)$
Division	$\frac{A}{B} - \frac{a\|B\| + b\|A\|}{B^2 + \text{sgn}(AB)b\|B\|}$	$\frac{A}{B} + \frac{a\|B\| + b\|A\|}{B^2 - \text{sgn}(AB)b\|B\|}$

NOTE.—A ± a and B ± b are error bounds for x_1 and x_2. The formulas give upper and lower error bounds for the four basic mathematical operations performed with x_1 and x_2. The formulas are valid only when x_1 and x_2 are functionally independent variables.

Let $y = x_1x_2$ where x_1 and x_2 are independent variables; let A and B represent the measured values of x_1 and x_2; let ε_A and ε_B signify the (exact) errors associated with A and B; and let a and b stand for upper bounds on $|\varepsilon_A|$ and $|\varepsilon_B|$. Then

$$y = x_1x_2 = (A + \varepsilon_A)(B + \varepsilon_B) \tag{B-35}$$

$$y = AB + \varepsilon_A B + \varepsilon_B A + \varepsilon_A\varepsilon_B \tag{B-36}$$

Letting $y = AB + \varepsilon_{AB}$, the (exact) error in the product ε_{AB} is given by

$$\varepsilon_{AB} = \varepsilon_A B + \varepsilon_B A + \varepsilon_A\varepsilon_B \tag{B-37}$$

Since in this case there is no statistical basis for a partial cancellation of errors, a worst-case analysis is the only recourse. Hence, the maximum and minimum values of ε_{AB} for $-a \leq \varepsilon_A \leq a$ and $-b \leq \varepsilon_B \leq b$ must be determined.

Assuming $a < |A|$ and $b < |B|$, the results are readily found to be as illustrated in Table A-7.

The four cases can be combined into the single formula $\text{sgn}(AB)ab \pm (a|B| + b|A|)$, where sgn(AB) denotes the algebraic sign of the product AB.

---

(20) Jenson, V. G., and G. V. Jeffreys. Mathematical Methods in Chemical Engineering. Academic Press, New York, New York, 1963, pp. 356-360.

TABLE B-7. MAXIMUM AND MINIMUM ERRORS IN THE PRODUCT AB

sgn A	sgn B	ε_{AB} Maximum	ε_{AB} Minimum
+	+	ab + aB + bA	ab - aB - bA
-	-	ab - aB - bA	ab + aB + bA
+	-	-ab - aB + bA	-ab + aB - bA
-	+	-ab + aB - bA	-ab - aB + bA

When a and b are small relative to |A| and |B|, the product ab can be neglected and the formula reduced to

$$|\varepsilon_{AB}|\max = a|B| + b|A| \tag{B-38}$$

Dividing by |AB| gives

$$\frac{|\varepsilon_{AB}|\max}{|AB|} = \frac{a}{|A|} + \frac{b}{|B|} \tag{B-39}$$

which states that the relative absolute errors are additive. This is the form given in Reference 20. It can be derived more easily by linearizing the problem at the outset. Thus, to terms of first order,

$$y = f(x_1, x_2) \cong f(A, B) + \frac{\partial f(A, B)}{\partial x_1}(x_1 - A) + \frac{\partial f(A, B)}{\partial x_2}(x_2 - B) \tag{B-40}$$

$$y - f(A, B) = \frac{\partial f(A, B)}{\partial x_1}(x_1 - A) + \frac{\partial f(A, B)}{\partial x_2}(x_2 - B) \tag{B-41}$$

$$\varepsilon_{AB} = \frac{\partial f(A, B)}{\partial x_1}\varepsilon_A + \frac{\partial f(A, B)}{\partial x_2}\varepsilon_B \tag{B-42}$$

Taking $f(x_1, x_2) = x_1 x_2$,

$$\varepsilon_{AB} = B\varepsilon_A + A\varepsilon_B \tag{B-43}$$

Thus,

$$(\varepsilon_{AB})_{\max} = |B|a + |A|b \tag{B-44}$$

$$(\varepsilon_{AB})_{\min} = -|B|a - |A|b \tag{B-45}$$

or

$$|\varepsilon_{AB}|_{\max} = a|B| + b|A| \tag{B-46}$$

From Equation B-42, the general formula for computing linearized systematic error bounds is

$$f(A, B) \pm \left[\left| \frac{\partial f(A, B)}{\partial x_1} \right| a + \left| \frac{\partial f(A, B)}{\partial x_2} \right| b \right]$$

For operations more complex than those given in Table B-6, error bounds are best determined by the above linearization procedure if errors are small or by direct substitution if errors are large. In the latter case, however, the bounds on y cannot generally be obtained by substituting the corresponding bounds on the x's into the function f. In fact, determining the bounds on y rigorously requires solving two nonlinear programming problems, namely,

$$\underset{\varepsilon_A, \varepsilon_B \cdot \cdot \cdot \varepsilon_N}{\text{Max}} \quad y = f\left(A + \varepsilon_A, B + \varepsilon_B, \ldots, N + \varepsilon_N\right) \qquad \text{(B-47)}$$

and

$$\underset{\varepsilon_A, \varepsilon_B \cdot \cdot \cdot \varepsilon_N}{\text{Min}} \quad y = f\left(A + \varepsilon_A, B + \varepsilon_B, \ldots, N + \varepsilon_N\right) \qquad \text{(B-48)}$$

subject to

$$-a \le \varepsilon_A \le a, -b \le \varepsilon_B \le b, \ldots, -n \le \varepsilon_N \le n \qquad \text{(B-49)}$$

Examples: Systematic Errors

Example 5--
For comparison, Example 1 is reworked assuming that the errors are systematic rather than random. Taking $A = 0.503$, $B = 0.316$, $a = 0.521$, and $b = 0.139$, the bounds on the sum are given by

$$A + B \pm (a + b) = 0.819 \pm 0.660$$

$$= 0.819 \pm 81\%$$

Of course, the systematic error bounds will always be larger than those for random errors because

$$\sqrt{a^2 + b^2} \le a + b \qquad \text{(B-50)}$$

according to the triangle inequality. A similar relationship can be demonstrated for the (linearized) multiplication and division formulas.

Example 6--
The specific reaction rate of a first order chemical reaction is $3.0 \times 10^{-7} \pm 50\%$ sec^{-1}, and the initial reactant concentration is $0.1 \pm 20\%$ kg-mole/m^3. It is desired to compute error bounds on the initial reaction rate $\gamma_o = kC_o$ assuming that the errors are systematic. For this problem, $A = 3 \times 10^{-7}$, $a = 1.5 \times 10^{-7}$, $B = 0.1$, and $b = 0.02$.

Using the multiplication formulas in Table B-6,

Lower bound--

$$\begin{aligned}(3 \times 10^{-8}) + (3 \times 10^{-9}) - (1.5 \times 10^{-8} + 6 \times 10^{-9}) &\\ &= (3.3 \times 10^{-8}) - (2.1 \times 10^{-8}) \\ &= 1.2 \times 10^{-8}\end{aligned}$$

Upper bound--

$$(3.3 \times 10^{-8}) + (2.1 \times 10^{-8}) = 5.4 \times 10^{-8}$$

Thus,

$$1.2 \times 10^{-8} \le (\gamma_o = 3 \times 10^{-8}) \le 5.4 \times 10^{-8}$$

or

$$\gamma_o = 3 \times 10^{-8} \begin{cases} + 2.4 \times 10^{-8} \\ - 1.8 \times 10^{-8} \end{cases} \text{kg-mole/m}^3\text{/s}$$

Using the linearized version of the formulas yields

$$0.9 \times 10^{-8} \le (\gamma_o = 3 \times 10^{-8}) \le 5.1 \times 10^{-8}$$

or

$$\gamma_o = 3 \times 10^{-8} \pm 2.1 \times 10^{-8} \text{ kg-mole/m}^3\text{/s}$$

This represents an error of 25% in the lower bound and 6% in the upper bound. However, the estimate for the lower bound is conservative, while that for the upper bound is not.

Unsymmetrical Systematic Errors

The formulas given in Table B-6 assume that the upper and lower error bounds are symmetrical about the nominal values of the variables. This is often an unrealistic assumption for systematic errors. For example, the algebraic sign of a systematic error may be known even though the magnitude can only be estimated. Unsymmetrical error bounds can also arise from physical restrictions on the variables; e.g., the variables may be restricted to non-negative values. In addition, use of the multiplication or division formulas in Table B-6 will result in unsymmetrical bounds. Thus, if a sequence of calculations is to be performed, formulas for variables having unsymmetrical error bounds are required.

Error propagation formulas are given in Table B-8 for the case in which

$$A - a_1 \leq x_1 \leq A + a_2 \qquad \text{(B-51)}$$

where $a_1, a_2 \geq 0$

and

$$B - b_1 \leq x_2 \leq B + b_2 \qquad \text{(B-52)}$$

where $b_1, b_2 \geq 0$

TABLE B-8. ERROR PROPAGATION FORMULAS FOR UNSYMMETRICAL SYSTEMATIC ERRORS

Operation	Lower bound	Upper bound
Addition	$A + B - (a_1 + b_1)$	$A + B + (a_2 + b_2)$
Subtraction	$A - B - (a_1 + b_2)$	$A - B + (a_2 + b_1)$
Multiplication	$AB + \text{sgn}(AB)\alpha_1\beta_1 - (\alpha_1\lvert B\rvert + \beta_1\lvert A\rvert)$	$AB + \text{sgn}(AB)\alpha_2\beta_2 + (\alpha_2\lvert B\rvert + \beta_2\lvert A\rvert)$
Division	$\frac{A}{B} - \frac{\alpha_1\lvert B\rvert + \beta_2\lvert A\rvert}{B^2 + \text{sgn}(AB)\beta_2\lvert B\rvert}$	$\frac{A}{B} + \frac{\alpha_2\lvert B\rvert + \beta_2\lvert A\rvert}{B^2 - \text{sgn}(AB)\beta_1\lvert B\rvert}$

where $\alpha_2 = a_1 \frac{1 + \text{sgn}B}{2} + a_2 \frac{1 - \text{sgn}B}{2}$

$\alpha_2 = a_1 \frac{1 + \text{sgn}B}{2} + a_2 \frac{1 + \text{sgn}B}{2}$

$\beta_1 = b_1 \frac{1 + \text{sgn}A}{2} + b_2 \frac{1 - \text{sgn}A}{2}$

$\beta_2 = b_1 \frac{1 - \text{sgn}A}{2} + b_2 \frac{1 + \text{sgn}A}{2}$

NOTE.—$A^{+a_2}_{-a_1}$ and $B^{+b_2}_{-b_1}$ are error bounds for x_1 and x_2. The formulas give upper and lower error bounds for the four basic mathematical operations performed with x_1 and x_2. The formulas are valid only when x_1 and x_2 are functionally independent variables.

These formulas can be derived in the same manner as those in Table B-6 although the multiplication and division formulas become somewhat cumbersome. The latter two sets of formulas are subject to the restrictions $a_1, a_2 < \lvert A\rvert$ and $b_1, b_2 < \lvert B\rvert$. It can be easily verified that the formulas in Table B-8 reduce to those in Table B-6 when $a_1 = a_2$ and $b_1 = b_2$.

Examples: Unsymmetrical Systematic Errors

Example 7--

In Example 5, $A = 0.503$, $a = 0.521$, $B = 0.316$, and $b = 0.139$. Thus, the lower bound on x_1 was $0.503 - 0.521 = -0.018$. However, x_1 and x_2 represent real emission factors and hence cannot be negative. Thus, it is more realistic to take $a_1 = 0.503$, $a_2 = 0.521$, and $b_1 = b_2 = 0.139$. Using the addition formula from Table B-8 yields

Lower bound--

$$A + B - (a_1 + b_1) = 0.819 - 0.642$$

$$= 0.177$$

Upper bound--

$$A + B + (a_2 + b_2) = 0.819 + 0.660$$

$$= 1.479$$

In Example 5, a lower bound of 0.159 was obtained for the sum. However, the present value of 0.177 is still more conservative than the value of 0.280 obtained in Example 1 using the formulas for random errors.

Example 8--

Obtain an error propagation formula for unsymmetrical systematic errors for a functional relationship of the form.

$$y = f(x_1, x_2) = \frac{x_1}{{x_2}^2} \qquad \text{(B-53)}$$

where $x_1, x_2 > 0$

Using the linearization procedure described previously, start with Equation B-42,

$$\varepsilon_{A/B^2} = \frac{\partial f(A, B)}{\partial x_1}\left(\varepsilon_A\right) + \frac{\partial f(A, B)}{\partial x_2}\left(\varepsilon_B\right) \qquad \text{(B-54)}$$

$$\varepsilon_{A/B^2} = \frac{1}{B^2}\left(\varepsilon_A\right) - \frac{2\,A}{B^3}\left(\varepsilon_B\right) \qquad \text{(B-55)}$$

where $-a_1 \leq \varepsilon_A \leq a_2$ and $-b_1 \leq \varepsilon_B \leq b_2$. It is assumed that the measured values, A and B, are both positive. Hence, it is readily seen that

$$\left(\varepsilon_{A/B^2}\right)_{\max} = \frac{a_2}{B^2} + \frac{2\,A}{B^3}(b_1) \qquad \text{(B-56)}$$

$$\left(\epsilon_{A/B^2}\right)_{min} = -\frac{a_1}{B^2} - \frac{2\ A}{B^3}(b_2) \qquad \text{(B-57)}$$

These equations can be expressed symbolically as follows:

$$\frac{A \begin{smallmatrix} +\ a_2 \\ \\ -\ a_1 \end{smallmatrix}}{\left(B \begin{smallmatrix} +\ b_2 \\ \\ -\ b_1 \end{smallmatrix}\right)^2} = \frac{A}{B^2} \begin{matrix} + \left[\dfrac{a_2}{B^2} + \dfrac{2\ A}{B^3}(b_1)\right] \\ \\ - \left[\dfrac{a_1}{B^2} + \dfrac{2\ A}{B^3}(b_2)\right] \end{matrix} \qquad \text{(B-58)}$$

Since both variables are positive, it is a simple matter to obtain the exact error bounds (as opposed to the linearized bounds) by direct substitution. Thus, the upper bound is obtained by substituting the largest value for the numerator and the smallest value for the denominator.

Upper bound--

$$\frac{A + a_1}{(B - b_1)^2} = \frac{A(1 + a_2/A)}{B^2(1 - b_1/B)^2} = \frac{A}{B^2}\left[1 + \frac{1 + a_2/A}{(1 - b_1/B)^2} - 1\right] \qquad \text{(B-59)}$$

$$= \frac{A}{B^2}\left[1 + \frac{(1 + a_2/A) - (1 - 2\ b_1/B + b_1{}^2/B^2)}{(1 - b_1/B)^2}\right] \qquad \text{(B-60)}$$

$$= \frac{A}{B^2}\left[1 + \frac{a_2/A + 2\ b_1/B - b_1{}^2/B^2}{(1 - b_1/B)^2}\right] \qquad \text{(B-61)}$$

Lower bound--

$$\frac{A - a_1}{(B + b_2)^2} = \frac{A(1 - a_1/A)}{B^2(1 + b_2/B)^2} = \frac{A}{B^2}\left[1 + \frac{1 - a_1/A}{(1 + b_2/B)^2} - 1\right] \qquad \text{(B-62)}$$

$$= \frac{A}{B^2}\left[1 + \frac{(1 - a_1/A) - (1 + 2\ b_2/B + b_2{}^2/B^2)}{(1 + b_2/B)^2}\right] \qquad \text{(B-63)}$$

$$= \frac{A}{B^2}\left[1 - \frac{(a_1/A + 2\ b_2/B + b_2{}^2/B^2)}{(1 + b_2/B)^2}\right] \qquad \text{(B-64)}$$

Thus,

$$\varepsilon_{A/B^2}\ \text{max} = \frac{A}{B^2}\left[\frac{a_2/A + 2\ b_1/B - b_1{}^2/B^2}{(1 - b_1/B)^2}\right] \tag{B-65}$$

$$\varepsilon_{A/B^2}\ \text{min} = \frac{-A}{B^2}\left[\frac{a_1/A + 2\ b_2/B + b_2{}^2/B^2}{(1 + b_2/B)^2}\right] \tag{B-66}$$

These results reduce to those obtained using the linearization procedure when the error in the denominator is small; i.e., when $b_{1/B} << 1$ and $b_{2/B} << 1$.

CONCLUSIONS

The assumptions underlying the error propagation formulas for random errors can be summarized as follows:

- The errors are normally distributed. (This assumption is required if normal statistics are used to compute confidence intervals.)
- The errors are statistically independent.
- When the error variances are estimated from the measured data, the formulas are valid for large sample sizes. In practice, a sample size of five should be adequate for most purposes.
- For nonlinear relations, the formulas are valid for small errors only; i.e., $a << |A|$ and $b << |B|$.

For systematic errors, the only assumptions are that the variables are functionally independent and, in the case of the multiplication and division formulas, $a < |A|$ and $b < |B|$. The linearized versions of these equations are accurate for small errors; i.e., $a << |A|$ and $b << |B|$.

When both types of errors are important in a calculation, they should be treated separately. The total error in the calculated value is then the sum of the random and systematic errors (15). When the type of error is not known, as is sometimes the case with data obtained from the literature, the formulas for systematic errors should be used since they yield more conservative error bounds. The systematic error formulas can also be used as a conservative approximation for propagation of random errors when the assumptions underlying the formulas for random errors are invalid. For the addition and subtraction of quantities subject to random errors, sharper bounds can be obtained by using the appropriate statistical method when the error variances are known to be equal.

APPENDIX C

DERIVATION OF SOURCE SEVERITY EQUATIONS

SUMMARY OF EQUATIONS

The source severity of pollutants can be calculated using the mass emission rate, Q, the effective height of the emissions, H, and the primary ambient air quality standard, PAAQS, or the threshold limit value, TLV. The severity equations shown in Table C-1 are derived in this appendix.[a]

TABLE C-1. SOURCE SEVERITY EQUATIONS FOR ELEVATED SOURCES

Pollutant	Severity equation
Particulate matter	$S = \frac{70\ Q}{H^2}$
SO_x	$S = \frac{50\ Q}{H^2}$
NO_x	$S = \frac{315\ Q}{H^{2.1}}$
Hydrocarbons	$S = \frac{162\ Q}{H^2}$
CO	$S = \frac{0.78\ Q}{H^2}$
Others	$S = \frac{5.5\ Q}{TLV(H^2)}$

DERIVATION OF EQUATION FOR $\overline{\chi}_{max}$

The most widely accepted formula for predicting time-averaged ground level concentrations downwind from a point source is the generalized Gaussian dispersion equation (5, 6).

---

[a] For convenience, the subscript "C" for "calculated" severity is omitted in this appendix.

$$\chi = \frac{Q}{\pi\sigma_y\sigma_z u} \exp\left[-\frac{1}{2}\left(\frac{y}{\sigma_y}\right)^2\right]\exp\left[-\frac{1}{2}\left(\frac{H}{\sigma_z}\right)^2\right] \qquad \text{(C-1)}$$

where χ = time-averaged ground level concentration at reference coordinates x and y, g/m^3

Q = mass emission rate, g/s

σ_y = lateral dispersion coefficient, m

σ_z = vertical dispersion coefficient, m

u = wind speed, m/s

y = horizontal distance from centerline of plume, m

H = effective height of emission release, m

x = downwind distance from source of emission release, m

$\pi \cong 3.14$

The maximum concentration occurs on the centerline of the plume where y = 0. Setting y = 0 in Equation C-1 yields

$$\chi = \frac{Q}{\pi_y\sigma_z\sigma u} \exp\left[-\frac{1}{2}\left(\frac{H}{\sigma_z}\right)^2\right] \qquad \text{(C-2)}$$

The downwind distance at which the maximum concentration occurs is found by solving the equation

$$\frac{d\chi}{dx} = 0 \qquad \text{(C-3)}$$

Differentiating Equation C-2 and substituting the derivative into Equation C-3 yields

$$\left(1 - \frac{H^2}{{\sigma_z}^2}\right)\frac{d\sigma_z}{dx} + \frac{\sigma_z}{\sigma_y}\left(\frac{d\sigma_y}{dx}\right) = 0 \qquad \text{(C-4)}$$

Equation C-4 can be solved if it is assumed that the lateral and vertical dispersion coefficients are identically equal; i.e.,

$$\sigma_y \equiv \sigma_z \qquad \text{(C-5)}$$

With this assumption, Equation B-4 becomes

$$\left(2 - \frac{H^2}{{\sigma_y}^2}\right)\frac{d\sigma_y}{dx} = 0 \qquad \text{(C-6)}$$

Since σ_y is a monotone increasing function of x, $d\sigma_y/dx \neq 0$. Therefore, Equation C-6 reduces to

$$2 - \frac{H^2}{{\sigma_y}^2} = 0 \qquad \text{(C-7)}$$

or

$$\sigma_y = \sigma_z = \frac{H}{\sqrt{2}} \qquad \text{(C-8)}$$

Substituting this result back into Equation C-2 yields the desired result,

$$\chi_{max} = \frac{2\ Q}{\pi e u H^2} \qquad \text{(C-9)}$$

The validity of the assumption $\sigma_y \equiv \sigma_z$ can be ascertained by referring to Tables C-2 and C-3 (21-24). These tables give the functional dependence of σ_y and σ_z on downwind distance. It can be seen that for stability Class C, Equation C-5 is approximately satisfied.

The averaging (i.e., sampling) time associated with Equation C-9 is approximately 3 min.[a] For averaging times between 3 min and 24 hr, a semiempirical correction factor given by Turner (5) is applied to Equation C-9.[b]

---

[a] The two sets of values in Tables C-2 and C-3 are not entirely consistent with respect to averaging time. The values for σ_y in Table C-2 correspond to an averaging time of 3 min. The values for σ_z in Table C-3 correspond to averaging times in excess of some limiting value, τ_m, which is approximately proportional to emission height for emission heights up to 100 m, at which height and above τ_m is approximately 10 min (24). Thus, Turner (5) states that concentrations calculated using these values for σ_y and σ_z correspond to an averaging time between 3 min and 10 min.

[b] The correction factor is applicable to situations in which the mean wind direction remains constant during the period of interest. That is, as the averaging time increases, the width of the plume increases, but the position of the plume centerline does not change. Since the mean wind direction does not, in general, remain constant over an extended period of time (e.g., 24 hr), this correction factor corresponds to a worst case situation with respect to pollutant concentration.

- - - - - - - - - - - -

(21) Eimutis, E. C., and M. G. Konicek. Derivations of Continuous Functions for the Lateral and Vertical Atmospheric Dispersion Coefficients. Atmospheric Environment, 6(11): 859-863, 1972.

(22) Tadmor, J., and Y. Gur. Analytical Expressions for the

(continued)

TABLE C-2. VALUES OF a FOR THE COMPUTATION OF σ_y[a] (21)

Stability class	a
A	0.3658
B	0.2751
C	0.2089
D	0.1471
E	0.1046
F	0.0722

[a] For the equation

$$\sigma_y = ax^b$$

where x = downwind distance
b = 0.9031 (from Reference 40)

TABLE C-3. VALUES OF THE CONSTANTS USED TO ESTIMATE VERTICAL DISPERSION[a] (23)

Usable range, m	Stability class	Coefficient		
		c_1	d_1	f_1
>1,000	A	0.00024	2.094	-9.6
	B	0.055	1.098	2.0
	C	0.113	0.911	0.0
	D	1.26	0.516	-13
	E	6.73	0.305	-34
	F	18.05	0.18	-48.6
		c_2	d_2	f_2
100 to 1,000	A	0.0015	1.941	9.27
	B	0.028	1.149	3.3
	C	0.113	0.911	0.0
	D	0.222	0.725	-1.7
	E	0.211	0.678	-1.3
	F	0.086	0.74	-0.35
		c_3	d_3	f_3
<100	A	0.192	0.936	0
	B	0.156	0.922	0
	C	0.116	0.905	0
	D	0.079	0.881	0
	E	0.063	0.871	0
	F	0.053	0.814	0

[a] For the equation

$$\sigma_z = cx^d + f$$

$$\bar{\chi}_{max} = \chi_{max}\left(\frac{t_o}{t}\right)^{0.17} \qquad \text{(C-10)}$$

or

$$\bar{\chi}_{max} = \frac{2\ Q}{\pi e u H^2}\left(\frac{t_o}{t}\right)^{0.17} \qquad \text{(C-11)}$$

where t_o = "short-term" averaging time; i.e., 3 min

t = averaging time of interest

$\bar{\chi}_{max}$ = maximum ground level concentration corresponding to averaging time, t

DEVELOPMENT OF SOURCE SEVERITY EQUATIONS

Source severity, S, has been defined as follows:

$$S = \frac{\bar{\chi}_{max}}{F} \qquad \text{(C-12)}$$

where $\bar{\chi}_{max}$ is given by Equation C-11 with u = 4.5 m/s

and where F = PAAQS for criteria pollutants
= TLV/300 for noncriteria pollutants

The averaging time associated with $\bar{\chi}_{max}$ is 24 hr for noncriteria pollutants; for criteria pollutants, it is the value specified in the corresponding PAAQS, as shown in Table C-4 (25).

(continued)
Vertical and Lateral Dispersion Coefficients in Atmospheric Diffusion. Atmospheric Environment, 3:688-689, 1969.

(23) Martin, D. O., and J. A. Tikvart. A General Atmospheric Diffusion Model for Estimating the Effects on Air Quality of One or More Sources. Presented at the 61st Annual Meeting of the Air Pollution Control Association, St. Paul, Minnesota, June 23-27, 1968. 18 pp.

(24) Pasquill, F. Atmospheric Dispersion Parameters in Gaussian Plume Modelling, Part II, Possible Requirements for Change in the Turner Workbook Values. EPA-600/4-76-030b (PB 258 036), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, June 1976. 43 pp.

(25) Code of Federal Regulations, Title 42--Public Health, Chapter IV--Environmental Protection Agency, Part 410--National Primary and Secondary Ambient Air Quality Standards, April 28, 1971. 16 pp.

TABLE C-4. SUMMARY OF NATIONAL AMBIENT AIR QUALITY STANDARDS (25)

Pollutant	Averaging time, hr	Primary standards	Secondary standards
		$\mu g/m^3$ (ppm)	
Particulate	Annual (geometric mean)	75	60[a]
	24[b]	260	150
SO_x	Annual	80 (0.03)	60 (0.02)
	(arithmetic mean)	24	365 (0.14)
	24[b]	365 (0.14)	260[c] (0.1)
	3[b]	None.	1,300 (0.5)
CO	8[b]	10,000 (9)	None.
	1[b]	40,000 (35)	40,000 (35)
Nitrogen dioxide	Annual (arithmetic mean)	100 (0.05)	100 (0.05)
Photochemical oxidants	1[b]	160 (0.08)	160 (0.08)
Hydrocarbons (nonmethane)[d]	3 (6 a.m. to 9 a.m.)	160 (0.24)	160 (0.24)

[a] The secondary annual standard (60 $\mu g/m^3$) is a guide for assessing implementation plans to achieve the 24-hr secondary standard.

[b] Not to be exceeded more than once per year.

[c] The secondary annual standard (260 $\mu g/m^3$) is a guide for assessing implementation plans to achieve the annual standard.

[d] Recommended guideline for meeting the primary ambient air quality standard for photochemical oxidants.

CO Severity

The primary standard for CO is reported for a 1-hr averaging time. Therefore,

$$t = 60 \text{ min}$$
$$t_o = 3 \text{ min}$$

$$\bar{X}_{max} = X_{max}\left(\frac{3}{60}\right)^{0.17} \qquad \text{(C-13)}$$

$$= \frac{2\,Q}{\pi e u H^2}\left(\frac{3}{60}\right)^{0.17} \qquad \text{(C-14)}$$

$$= \frac{2\,Q}{(3.14)(2.72)(4.5)H^2}(0.6) \qquad \text{(C-15)}$$

$$\overline{\chi}_{max} = \frac{0.052\ Q}{H^2}(0.6) \qquad \text{(C-16)}$$

$$\overline{\chi}_{max} = \frac{(3.12 \times 10^{-2})Q}{H^2} \qquad \text{(C-17)}$$

Setting F equal to the primary standard for CO, i.e., 0.04 g/m^3, yields

$$S = \frac{\overline{\chi}_{max}}{F} = \frac{(3.12 \times 10^{-2})Q}{0.04\ H^2} \qquad \text{(C-18)}$$

or

$$S \cong \frac{0.78\ Q}{H^2} \qquad \text{(C-19)}$$

Hydrocarbon Severity

For nonmethane hydrocarbons, a 3-hr averaging time is used.

$$t = 180 \text{ min}$$

$$\overline{\chi}_{max} = \chi_{max}\left(\frac{3}{180}\right)^{0.17} \qquad \text{(C-20)}$$

$$= 0.5\ \chi_{max} \qquad \text{(C-21)}$$

$$= \frac{(0.5)(0.052)Q}{H^2} \qquad \text{(C-22)}$$

$$\overline{\chi}_{max} = \frac{0.026\ Q}{H^2} \qquad \text{(C-23)}$$

For nonmethane hydrocarbons, the concentration of 1.6 x 10^{-4} g/m^3 has been established as a guide for achieving oxidant standards. Therefore,

$$S = \frac{0.026\ Q}{(1.6 \times 10^{-4})H^2} \qquad \text{(C-24)}$$

or

$$S \cong \frac{162.5\ Q}{H^2} \qquad \text{(C-25)}$$

Particulate Severity

The primary standard for particulate matter is reported for a 24-hr averaging time. Therefore,

$$\overline{\chi}_{max} = \chi_{max}\left(\frac{3}{1,440}\right)^{0.17} \qquad (C-26)$$

$$= \frac{(0.052)Q(0.35)}{H^2} \qquad (C-27)$$

$$\overline{\chi}_{max} = \frac{0.0182\ Q}{H^2} \qquad (C-28)$$

The primary standard for particulate matter is 2.6×10^{-4} g/m^3. Thus,

$$S = \frac{0.0182\ Q}{(2.6 \times 10^{-4})H^2} \qquad (C-29)$$

or

$$S \cong \frac{70\ Q}{H^2} \qquad (C-30)$$

SO_x Severity

The primary standard for SO_x is 3.65×10^{-4} g/m^3 for an averaging time of 24 hr. Thus, proceeding as before,

$$\overline{\chi}_{max} = \frac{0.0182\ Q}{H^2} \qquad (C-31)$$

$$S = \frac{\overline{\chi}_{max}}{F} = \frac{0.0182\ Q}{(3.65 \times 10^{-4})H^2} \qquad (C-32)$$

or

$$S \cong \frac{50\ Q}{H^2} \qquad (C-33)$$

NO_2 Severity

Since NO_2 has a primary standard with a 1-yr averaging time, Equation C-10 cannot be used to calculate χ_{max}. Hence, the following equation is used which gives the annual mean ground level concentration (2):[a]

$$\overline{\chi} = \frac{2.03\ Q}{\sigma_z u x} \exp\left[-\frac{1}{2}\left(\frac{H}{\sigma_z}\right)\right]^2 \qquad \text{(C-34)}$$

To obtain the equation for $\overline{\chi}_{max}$, Equation C-34 is differentiated with respect to x, and the derivative is substituted into Equation C-3. These operations result in the following equation:

$$x\left(\frac{d\sigma_z}{dx}\right)\left(1 - \frac{H^2}{\sigma_z^{\ 2}}\right) + \sigma_z = 0 \qquad \text{(C-35)}$$

From Table C-3, for stability Class C, σ_z has the form

$$\sigma_z = ax^b \qquad \text{(C-36)}$$

Substituting Equation C-36 into C-35 and solving for x yields

$$x_{max} = \left(\frac{H}{a\sqrt{\frac{b+1}{b}}}\right)^{1/b} \qquad \text{(C-37)}$$

Using the values a = 0.113 and b = 0.911 from Table C-3 yields

$$x_{max} = \frac{H^{1.098}}{0.137} \qquad \text{(C-38)}$$

and

$$\left(\sigma_z\right)_{max} = a\ x^b_{\ max} = 0.691\ H \qquad \text{(C-39)}$$

Substituting these values for x and σ_z back into Equation C-34 and setting u = 4.5 m/s gives

$$\overline{\chi}_{max} = \frac{0.0314\ Q}{H^{2.098}} \qquad \text{(C-40)}$$

---

[a] Equation C-34 is based on a 16-point wind rose; i.e., the 360° of the compass are divided into 16 sectors of 22.5° each. Equation C-34 represents a worst case situation in that it assumes that the wind always blows from the same sector during the entire year. Thus, for example, if the wind directions were distributed equally over the 16 sectors during the year, the corresponding concentration would be one-sixteenth of the value given by Equation C-34.

Since the NO_2 standard is 1.0×10^{-4} g/m³, the NO_2 severity equation is

$$S = \frac{0.0314\ Q}{(1.0 \times 10^{-4}) H^{2.098}} \tag{C-41}$$

or

$$S \cong \frac{315\ Q}{H^{2.1}} \tag{C-42}$$

Noncriteria Emissions

For noncriteria pollutant concentrations, the averaging time is 24 hr. Thus,

$$\bar{\chi}_{max} = \chi_{max}\left(\frac{3}{1{,}440}\right)^{0.17} \tag{C-43}$$

$$= \frac{0.052\ Q}{H^2}(0.35) \tag{C-44}$$

$$\bar{\chi}_{max} = \frac{0.0182\ Q}{H^2} \tag{C-45}$$

Since F = TLV/300, the equation for source severity is

$$S = \frac{\bar{\chi}_{max}}{F} = \frac{0.0182\ Q}{(TLV/300) H^2} \tag{C-46}$$

or

$$S \cong \frac{5.5\ Q}{TLV(H^2)} \tag{C-47}$$

AFFECTED POPULATION CALCULATION

The affected population is calculated using Equation C-34 to predict average concentration, $\bar{\chi}$, as a function of downwind distance. The value of σ_z corresponding to stability Class C is used in Equation C-34; i.e.,

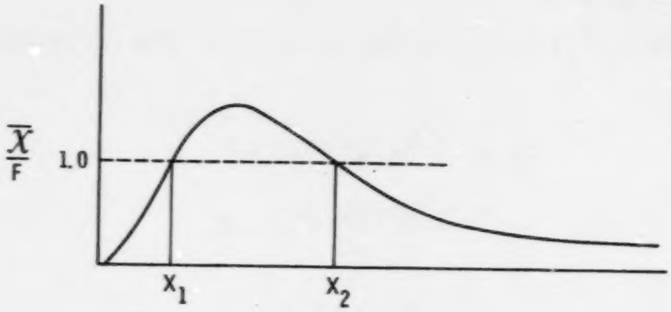
$$\sigma_z = 0.113\ x^{0.911} \tag{C-48}$$

The downwind distances at which $\overline{\chi}/F = 1.0$ are determined by an iterative technique;[a] i.e., the roots of the following equation are determined:

$$\frac{2.03\ Q}{\sigma_z u x F} \exp\left[-\frac{1}{2}\left(\frac{H}{\sigma_z}\right)^2\right] - 1.0 = 0 \tag{C-49}$$

where σ_z is given by Equation C-48.

A typical plot of $\overline{\chi}/F$ as a function of downwind distance appears as follows:



where x_1 and x_2 are the roots of Equation C-49

As previously noted, Equation C-34 gives the concentration for a worst-case situation in which the wind blows continuously from one sector of width 22.5° (one-sixteenth of a circle). Hence, the area in that sector over which $\overline{\chi}/F$ exceeds 1.0 is given by (see following sketch).

$$A' = \pi(x_2{}^2 - x_1{}^2) \tag{C-50}$$

The population density, D_P, in the vicinity of the representative source is obtained by averaging the county population densities for each plant in the source type. The worst case arises when most of the population is concentrated in the sector downwind from the source; e.g., in a metropolitan area.[b] This situation is approximated by assuming that the entire population in the vicinity of the representative source is distributed uniformly in the downwind sector.

---

[a] This calculation is not entirely consistent in that different averaging times are used for $\overline{\chi}$ and F. Equation C-34 gives an annual or long-term average value for $\overline{\chi}$. The averaging times for F vary from 1 hr to 1 yr. This discrepancy is considered acceptable in view of the very crude overall approach used to estimate the exposed population.

[b] The worst case arises when the wind blows from one sector most of the time and most of the population is contained in the downwind sector. The present approximation can be considered as a limiting case in which the wind always blows from the same sector and the entire population is contained in the downwind sector.

The population, P, in the area over which $\overline{\chi}/F>1.0$ is then given by

$$P = D_p(x_2^2 - x_1^2) \tag{C-51}$$

or

$$P = D_pA \text{ (persons)} \tag{C-52}$$

where

$$A = \pi(x_2^2 - x_1^2) \tag{C-53}$$

is the area contained in an annular region surrounding the source. The quantity, P, is designated the effected population.

APPENDIX D

PLUME RISE

The problem of determining plume rise from an elevated point source is discussed in this appendix. Equations are recommended for calculating plume rise in the Source Assessment Program, and error bounds for the calculated plume rise are estimated. These error bounds are employed in Section 5 and Appendix I.

RECOMMENDED PROCEDURE

According to Briggs (26), as of 1969 there were over 30 plume rise formulas in the literature with new ones appearing at a rate of about two per year. This multitude of formulas can be classified into two broad categories: 1) those which give the plume rise as a function of downwind distance and 2) those which yield a single value for the plume rise. Although the former equations show better agreement with observed plume rise, they are incompatible with the simple Gaussian dispersion equation used in Source Assessment. Of the equations in the second category, a modified Holland equation appears (on the basis of extensive data presented in References 26 and 27) to be about as good as any. The original Holland equation is (27)

$$\Delta H = \left(\frac{1}{u}\, 1.5\, V_s D + 0.04\, Q_H\right) \qquad \text{(D-1)}$$

where ΔH = plume rise, m

u = wind speed, m/s

V_s = stack gas exit velocity, m/s

D = stack diameter, m

Q_H = heat emission rate, kcal/s

---

(26) Briggs, G. A. Plume Rise. AEC Critical Review Series. U.S. Atomic Energy Commission, Division of Technical Information Extension, Oak Ridge, Tennessee, 1969. 81 pp.

(27) Moses, H., and M. R. Kraimer. Plume Rise Determination--A New Technique Without Equations. Journal of the Air Pollution Control Association, 22(8):621, 1972.

The heat emission rate is calculated as follows (30):

$$Q_H = \left(\frac{\pi D^2}{4}\right) V_s \left(\frac{PM}{R}\right) C_P \left(\frac{T_s - T_a}{T_s}\right) \qquad \text{(D-2)}$$

where
P = atmospheric pressure, dyne/m^2
R = gas constant = 8.314×10^5 dyne-m/gmole-°K
M = molecular weight of effluent, g/gmole
C_P = heat capacity at constant pressure of effluent, kcal/g-°K
T_s = stack gas temperature, °K
T_a = ambient air temperature, °K

The values calculated from Equation D-1 are generally too low. For example, values of $\Delta H_{calc}/\Delta H_{observed}$ for 22 stacks given in Reference 26 range from a low of 0.04 to a high of 1.18. Excluding the lowest and highest values, the range is 0.18 to 0.66. The mean is 0.45. The stack heights range from 60 m to 180 m.

Extensive data are also summarized in Reference 30. For small stacks (30 m to 40 m), the average value of $\Delta H_{calc}/\Delta H_{obs}$ is 0.17. For medium-size stacks (60 m to 120 m), the average value is 0.43. For the largest stacks (150 m to 180 m), the average is 1.0. For another set of data (Bringfelt data from Sweden) covering 36 stacks ranging from 20 m to 150 m in height, the average value of $\Delta H_{calc}/\Delta H_{obs}$ is 0.5. The average for all four sets of data in Reference 27 is 0.33.

Based on the value of 0.33 for $\Delta H_{calc}/\Delta H_{obs}$, the plume rise calculated from Holland's equation should be multiplied by a factor of 3. This value is in good agreement with a value of 2.92 recommended by Stumke (26) based on independent observations. A correction factor of 2.2 results from the data given by Briggs (29) (2.2 = 1/0.45).

For electric generating stations, an equation is available which has been optimized for best fit to existing data (28).

$$\Delta H = 0.414\left(\frac{1{,}000\ Q_H}{u^{0.694}}\right)^{0.444} = (8.892)\frac{Q_H{}^{0.444}}{u^{0.694}} \qquad \text{(D-3)}$$

---

(28) Thomas, F. W., S. B. Carpenter, and W. C. Colbaugh. Plume Rise estimates for electric Generating Stations. Journal of the Air Pollution Control Association, 20(3):170, 1970.

The following plume rise equations are recommended for source assessment:

For general sources,

$$\Delta H = \left(\frac{3}{u}\ 1.5\ V_s D + 0.04\ Q_H\right) \qquad \text{(D-4)}$$

For electric utility combustion sources,

$$\Delta H = (8.892)\frac{{Q_H}^{0.444}}{u^{0.694}} \qquad \text{(D-3)}$$

ESTIMATED ERROR BOUNDS FOR PLUME RISE

Based on data in References 26 and 27, error bounds for plume rise when using Equation D-4 are estimated to be

$$0.5 \leq \frac{\Delta H_{calc}}{\Delta H_{true}} \leq 3.0$$

$$0.33 \leq \frac{\Delta H_{true}}{\Delta H_{calc}} \leq 2.0$$

$$0.33\ \Delta H_{calc} \leq \Delta H_{true} \leq 2\ \Delta H_{calc}$$

$$\Delta H_{calc}(1.0 - 0.67) \leq \Delta H_{true} \leq \Delta H_{calc}(1.0 + 1.0)$$

Therefore,

$$\frac{c_{su}}{\Delta H_{calc}} = 1.0; \quad \frac{c_{s\ell}}{\Delta H_{calc}} = 0.67$$

For the error analysis, values of $\hat{c}_{su}$ and $\hat{c}_{s\ell}$ are required where

$$\hat{c}_{su} = \frac{c_{su}}{C} = \frac{c_{su}}{H_s + \Delta H_{calc}} \qquad \text{(D-5)}$$

and

$$\hat{c}_{s\ell} = \frac{c_{s\ell}}{C} = \frac{c_{s\ell}}{H_s + \Delta H_{calc}} \qquad \text{(D-6)}$$

Since

$$\hat{c}_{su} = \frac{c_{su}}{\Delta H_{calc}}\left(\frac{\Delta H_{calc}}{H_s + \Delta H_{calc}}\right) \tag{D-7}$$

values of the term in parenthesis are needed. The following data were obtained from Reference 27:

Stack size	Plant	$\Delta H_{true}/H_s$	
Small	Argonne I	0.17	Avg = 0.2
Small	Argonne II	0.25	
Medium	Harwell	0.87	Avg = 0.5
Medium	Gernsheim	0.38	
Medium	Duisberg	0.31	
Large	Lakeview	1.34	Avg = 1.2
Large	Widows Creek	1.14	
Large	Gallatin	1.49	
Large	Paradise	0.89	

For small stacks,

$$\Delta H_{true} = 0.2\ H_s \text{ and } \frac{\Delta H_{calc}}{\Delta H_{true}} = 0.5 \longrightarrow \Delta H_{calc} = 0.1\ H_s$$

Therefore,

$$\frac{\Delta H_{calc}}{H_s + \Delta H_{calc}} = \frac{0.1}{1.1} = 0.09$$

For large stacks,

$$\Delta H_{true} = 1.2\ H_s \text{ and } \frac{\Delta H_{calc}}{\Delta H_{true}} = 3.0 \rightarrow \Delta H_{calc} = 3.6\ H_s$$

Therefore,

$$\frac{\Delta H_{calc}}{H_s + \Delta H_{calc}} = \frac{3.6}{4.6} = 0.78$$

Note that c_{su} is a result of errors for small stacks and $c_{s\ell}$ for large stacks. Thus, error bounds for plume rise calculated from Equation D-4 are

$$\hat{c}_{su} = 1.0(0.09) \cong 0.1 \tag{D-8}$$

$$\hat{c}_{s\ell} = 0.67(0.78) \cong 0.5 \qquad \text{(D-9)}$$

Error bounds for plume rise calculated from Equation D-3 are estimated from data given in Reference 28 as follows:

$$0.6\ \Delta H_{calc} \leq \Delta H_{true} \leq 1.4\ \Delta H_{calc}$$

Therefore,

$$\frac{c_{su}}{\Delta H_{calc}} = 0.4;\ \frac{c_{s\ell}}{\Delta H_{calc}} = 0.4$$

The average stack height for data on which Equation D-3 is based is 130 m (range is 76 m to 183 m). Therefore, assume $\Delta H_{true} = 0.85\ H_s$, the average for medium and large stacks.

Upper Bound

$$\frac{\Delta H_{calc}}{\Delta H_{true}} = \frac{1}{1.4} = 0.71$$

$$\Delta H_{calc} = 0.71\ \Delta H_{true} = (0.71)0.85\ H_s = 0.61\ H_s$$

$$\frac{\Delta H_{calc}}{H_s + \Delta H_{calc}} = \frac{0.61}{1.61} = 0.38$$

$$\hat{c}_{su} = \frac{c_{su}}{\Delta H_{calc}}\left(\frac{\Delta H_{calc}}{H_s + \Delta H_{calc}}\right) = 0.4(0.38) = 0.15$$

Lower Bound

$$\frac{\Delta H_{calc}}{\Delta H_{true}} = \frac{1}{0.6} = 1.67$$

$$\Delta H_{calc} = (1.67)0.85\ H_s = 1.42\ H_s$$

$$\frac{\Delta H_{calc}}{H_s + \Delta H_{calc}} = \frac{1.42}{2.42} = 0.59$$

$$\hat{c}_{s\ell} = 0.4(0.59) = 0.23$$

Thus, error bounds for plume rise calculated from Equation D-3 are

$$\hat{c}_{su} = 0.15$$

$$\hat{c}_{s\ell} = 0.23$$

APPENDIX E

ALTERNATIVE METHODS FOR ESTIMATING "ACCEPTABLE" CONCENTRATION FOR NONCRITERIA POLLUTANTS

METHOD ORIGINALLY ADOPTED FOR SOURCE ASSESSMENT PROGRAM

For noncriteria pollutants, the "acceptable" concentration, F, is specified as follows:

$$F = \frac{\text{TLV}}{300} = (\text{TLV})\left(\frac{8}{24}\right)\left(\frac{1}{100}\right) \qquad \text{(E-1)}$$

The conversion factor, G = 1/300, converts the TLV to an "equivalent" primary ambient air quality standard. The factor 8/24 adjusts the TLV from an 8-hr work day to continuous (24-hr) exposure, and the factor of 1/100 is designed to account for the fact that the general population constitutes a higher risk group than healthy workers. This latter factor was chosen after consultation with EPA health effects experts.

METHOD BASED ON CRITERIA POLLUTANTS

An alternative method for estimating "acceptable" concentration for noncriteria pollutants is to determine the relationship between TLV's and primary ambient air quality standards (PAAQS) based on the criteria pollutants. This relationship is shown in Table E-1. The (geometric) mean conversion factor (i.e., $\text{PAAQS}_{24}/\text{TLV}$) for the four pollutants listed in the table is 0.0467. Thus,

$$F = 0.0467 \text{ TLV} = \frac{\text{TLV}}{21.4} \qquad \text{(E-2)}$$

The "acceptable" concentration based on data for criteria pollutants is therefore higher than that adopted for the Source Assessment Program by a factor of about 15.

Assuming that criteria pollutants constitute a random sample of size 4 from the population of all pollutants, and further assuming that the conversion factors are log-normally distributed over this population, a 95% confidence interval for the geometric mean conversion factor is found to be (0.0123, 0.179). The corresponding confidence interval for the inverse of the mean conversion factor is (5.6, 81.3). Thus, the value of 300 obtained by

TABLE E-1. RELATIONSHIP BETWEEN TLV'S AND PRIMARY AMBIENT AIR QUALITY STANDARDS

Criteria pollutant	PAAQS, $\mu g/m^3$	Averaging time, (t_{avg}), hr	Estimated 24-hr $PAAQS_{24}$,[a] $\mu g/m^3$	TLV, $\mu g/m^3$	Conversion factor, $PAAQS_{24}/TLV$
Particulate matter	260	24	260	10,000	0.0260
Sulfur dioxide	365	24	365	13,000	0.0281
Carbon monoxide	10,000	8	8,670	55,000	0.158
Nitrogen dioxide	100	8,760	370	9,000	0.0411

[a] Calculated using the method given in Reference 29. The equation used was

$$PAAQS_{24} = PAAQS(t_{avg}/24)^b$$

where b = 0.13 for CO and 0.22 for NO_2. These exponents are averages of the values given in Table 2 of Reference 32.

Method 1 lies well beyond the upper end point, 81.3, of this interval.

METHOD BASED ON LD_{50} DATA

The problem of estimating permissible concentrations of pollutants from available health effects data has been studied by Handy and Schindler (30). Their results yield the following two estimates of "acceptable" concentration in terms of animal LD_{50} values:[a]

$$F = 1.07 \times 10^{-4} \ (LD_{50}) \qquad \text{(E-3)}$$

$$F = 8.1 \times 10^{-5} \ (LD_{50}) \qquad \text{(E-4)}$$

Equation E-3 is based on a statistical correlation between TLV's and animal LD_{50} values for some 240 different compounds (30).

---

[a] LD_{50} is the dosage which results in mortality to 50% of exposed population.

- - - - - - - - - - -

(29) McGuire, T., and K. C. Knoll. Relationship Between Concentrations of Atmospheric Pollutants and Averaging Times. Atmospheric Environment, 5(5):291, 1971.

(30) Handy, R., and A. Schindler. Estimation of Permissible Concentrations of Pollutants for Continuous Exposure. EPA-600/2-76-155 (PB 253 959), U.S. Environmental Protection Agency, Washington, D.C., June 1976. 136 pp.

Equation E-4 is based on an assumed permissible body concentration of pollutant together with an assumed biological half-life of 30 days (30). Thus, nearly identical results are obtained by two entirely different approaches.

In order to compare Equations E-3 and E-4 with Equations E-1 and E-2, the least squares fit to the relationship between TLV and LD_{50} data obtained in Reference 30 is utilized:

$$LD_{50} = 34.5 \text{ (TLV)} \tag{E-5}$$

With this relationship, Equations E-3 and E-4 become

$$F = 0.0037 \text{ (TLV)} \tag{E-6}$$

$$F = 0.0028 \text{ (TLV)} \tag{E-7}$$

For comparison, Equation E-1 is rewritten as follows:

$$F = 0.0033 \text{ (TLV)} \tag{E-8}$$

Clearly, Equations E-1, E-3, and E-4 are essentially equivalent definitions for F. In fact, the coefficient 0.0033 in Equation E-8 is the mean of the two coefficients in Equations E-6 and E-7. It should be noted, however, that for a particular compound, the value of F based on TLV may differ by more than an order of magnitude from the value based on LD_{50}. This variability is due to the large deviations of individual data points from the least squares fit to the TLV versus LD_{50} data (30).

METHOD BASED ON DOSE-RESPONSE RELATIONSHIPS

The three previous methods have dealt with the estimation of an "acceptable" concentration using currently available health effects data. The present subsection considers the hypothetical situation in which a dose-response relationship for the human population has been established for a particular pollutant. Such a relationship is illustrated in Figure E-1.

In this figure, probability, P, of a lifetime response to pollutant exposure is given as a function of the concentration, C, to which exposure occurs. The solid line represents the curve fitted to experimental data obtained at high concentrations. The dotted line represents a statistical extrapolation (31) of

---

(31) Cornfield, J. Carcinogenic Risk Assessment. Science, 198(4318):693-699, 1977.

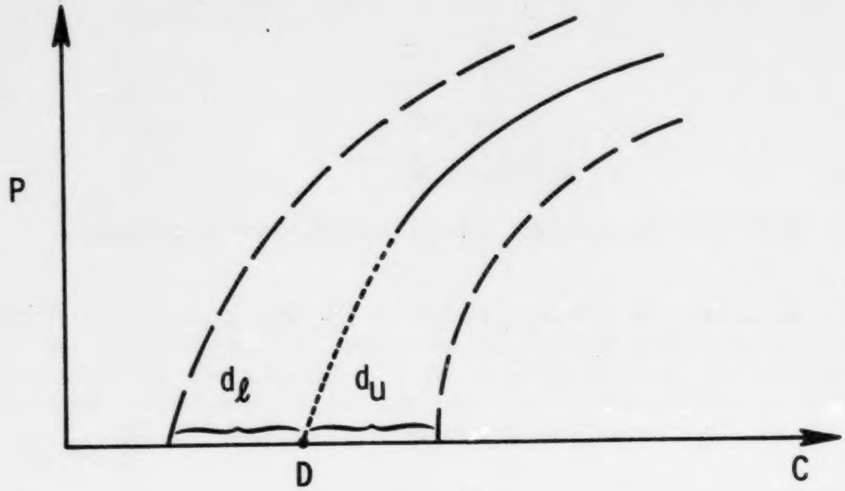


Figure E-1. Schematic representation of a dose-response relationship, with a positive no-response level.

the experimental curve to the no-response level (P = 0). The dashed lines represent 95% confidence limits for the experimental curve.

The estimated value, D, of "acceptable" concentration in this case would be the concentration at which the extrapolated experimental curve intersects the C-axis. The associated uncertainties, d_ℓ and d_u, would be obtained from the 95% confidence limits as indicated in Figure E-1.

The positive no-response level exhibited by the dose-response relationship shown in Figure E-1 is not the only type of behavior that can occur at low concentrations. The extrapolated experimental curve may pass through the origin as illustrated in Figure E-2, or it may intersect the P-axis as illustrated in Figure E-3. In order to determine an "acceptable concentration in the latter two cases, an "acceptable" probability, P*, of a response to pollutant exposure would have to be arbitrarily chosen. The estimated value, D, of "acceptable" concentration and the assoiated uncertainties, d_ℓ and d_u, would then be obtained as shown in Figures E-2 and E-3.

APPENDIX F

CONSTRUCTION OF OPERATING CHARACTERISTIC CURVES

This appendix explains the construction of the operating characteristic (OC) curves presented in Section 5. Attention is focused specifically on construction of the curves presented in Figure 4. For convenience, this figure is reproduced here as Figure F-1, to which the plotted points have been added.

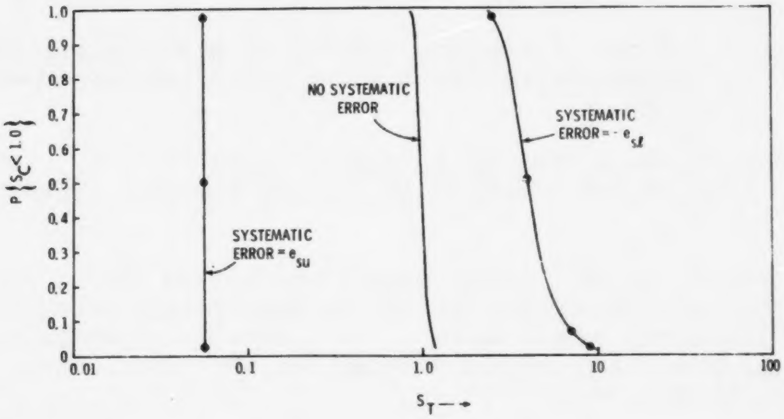


Figure F-1. Operating characteristic curves for Case A with $\hat{b}_r = 0.10$.

GOVERNING RELATIONSHIPS

From Equation 38 (Section 5) (taking $\hat{b}_r = 0.10$),

$$\hat{e}_u = \hat{e}_r + \hat{e}_{su} = 17.14 \qquad \text{(F-1)}$$

$$\hat{e}_\ell = \hat{e}_r + \hat{e}_{s\ell} = 0.89 \qquad \text{(F-2)}$$

$$\hat{e}_r = \sqrt{\hat{b}_r^{\,2} + 0.01} = 0.14 \qquad \text{(F-3)}$$

$$\hat{e}_{su} = 17 \qquad \text{(F-4)}$$

$$\hat{e}_{s\ell} = 0.75 \tag{F-5}$$

The uncertainty interval for source severity is therefore

$$S_C(1 - \hat{e}_{\ell}) \leq S_T \leq S_C(1 + \hat{e}_u) \tag{F-6}$$

or

$$0.11\, S_C \leq S_T \leq 18.14\, S_C \tag{F-7}$$

OC CURVE FOR SYSTEMATIC ERROR EQUAL TO $-e_{s\ell}$

The OC curve gives the (fiducial) probability that $S_T < 1.0$ as a function of S_C. The fiducial probability distribution of S_T is shown schematically in Figure F-2 for the case in which the systematic error is $-e_{s\ell} = -S_C\, \hat{e}_{s\ell}$. The distribution is centered at the point $S_C - e_{s\ell} = S_C(1 - \hat{e}_{s\ell})$. Since e_r represents a 95% confidence limit in this case, the 2.5% and 97.5% points of the distribution are located at a distance of $e_r = S_C\, \hat{e}_r$ on either side of the center point.

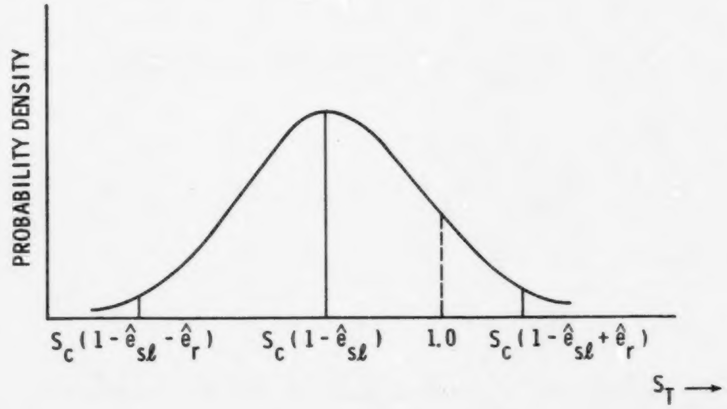

PROBABILITY DENSITY
$S_C(1-\hat{e}_{s\ell}-\hat{e}_r)$
$S_C(1-\hat{e}_{s\ell})$
1.0
$S_C(1-\hat{e}_{s\ell}+\hat{e}_r)$
$S_T \rightarrow$

Figure F-2. Schematic representation of fiducial distribution of S_T for systematic error equal to $-e_{s\ell}$.

From Figure F-2, when the central point of the distribution is at 1.0, $P\{S_T < 1.0\} = 0.5$. This situation occurs when

$$S_C(1 - \hat{e}_{s\ell}) = 1.0 \tag{F-8}$$

$$S_C = \frac{1}{1 - \hat{e}_{s\ell}} = \frac{1}{0.25} = 4.0 \tag{F-9}$$

The point $S_C = 4.0$, $P = 0.5$ is shown plotted in Figure F-1.

Similarly, the 0.975 probability point is found from the equation

$$S_C(1 - \hat{e}_{s\ell} + \hat{e}_r) = 1.0 \quad \text{(F-10)}$$

$$S_C = \frac{1}{0.25 + 0.14} = 2.56 \quad \text{(F-11)}$$

The point $S_C = 2.56$, $P = 0.975$ is also plotted in Figure F-1.

The 0.025% point is found in a similar manner.

$$S_C(1 - \hat{e}_{s\ell} - \hat{e}_r) = 1.0 \quad \text{(F-12)}$$

$$S_C = \frac{1}{0.25 - 0.14} = 9.1 \quad \text{(F-13)}$$

The point $S_C = 9.1$, $P = 0.025$ is shown plotted in Figure F-1.

Additional points on the OC curve are obtained by assuming that the probability distribution in Figure F-2 is a normal distribution.[a] For illustration, take $S_C = 7.0$. The mean of the distribution of S_T is then

$$S_C(1 - \hat{e}_{s\ell}) = 7.0(0.25) = 1.75 \quad \text{(F-14)}$$

The 97.5% point is

$$S_C(1 - \hat{e}_{s\ell} + \hat{e}_r) = 7.0(0.39) = 2.73 \quad \text{(F-15)}$$

The 2.5% point is

$$S_C(1 - \hat{e}_{s\ell} - \hat{e}_r) = 7.0(0.11) = 0.77 \quad \text{(F-16)}$$

The distribution is shown schematically in Figure F-3. The probability that $S_T < 1.0$ is equal to the area under the normal curve to the left of 1.0. In order to compute this area, the standard deviation of the distribution is needed. Since the 97.5% point of a normal distribution occurs at 1.96 standard deviations from the mean,

$$1.75 + 1.96\,\sigma = 2.73 \quad \text{(F-17)}$$

or

$$\sigma = 0.50 \quad \text{(F-18)}$$

---

[a] Note that the three points previously determined do not depend on this assumption. Thus, only the precise shape of the OC curve between these points is affected by this assumption. The precise shape of the curve is of little interest in the present study.

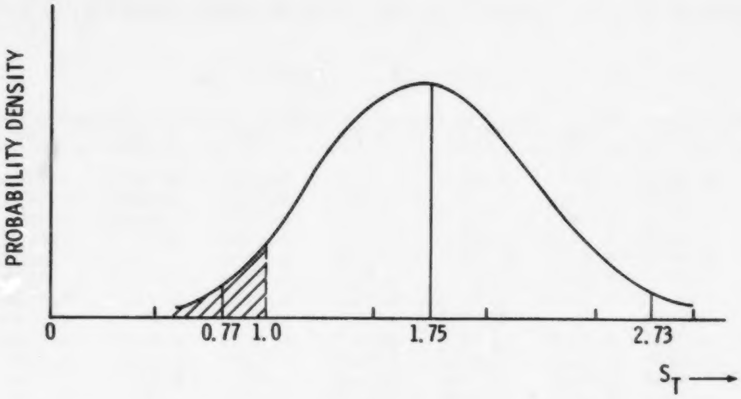


Figure F-3. Schematic representation of the distribution of S_T for systematic error equal to $-e_{s\ell}$ and $S_C = 7.0$.

Next, the standard normal deviate, z, corresponding to $S_T = 1.0$ is calculated.

$$z = \frac{S_T - \mu}{\sigma} \qquad \text{(F-19)}$$

$$z = \frac{1.0 - 1.75}{0.5} = -1.50 \qquad \text{(F-20)}$$

The area, F(z), under the standard normal curve to the left of z is given in Table F-1 (17). Because of the symmetry of the normal distribution,

$$F(-z) = 1 - F(z) \qquad \text{(F-21)}$$

Thus, entering Table F-1 with z = 1.50,

$$F(1.50) = 0.9332 \qquad \text{(F-22)}$$

$$F(-1.50) = 1 - F(1.50) = 0.0668 \qquad \text{(F-23)}$$

The point $S_C = 7.0$, $P = 0.0668$ is shown plotted in Figure F-1. The complete OC curve is constructed by repeating the above calculation for other values of S_C.

OC CURVE FOR SYSTEMATIC ERROR EQUAL TO e_{su}

The fiducial distribution of S_T is shown schematically in Figure F-4 for the case in which the systematic error is $e_{su} = S_C\hat{e}_{su}$. The 0.5 probability point is obtained from

$$S_C(1 + \hat{e}_{su}) = 1.0 \qquad \text{(F-24)}$$

$$S_C = \frac{1}{1 + 17} = 0.0556 \qquad \text{(F-25)}$$

TABLE F-1. AREA UNDER THE NORMAL CURVE, F(z) (14)

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9356	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

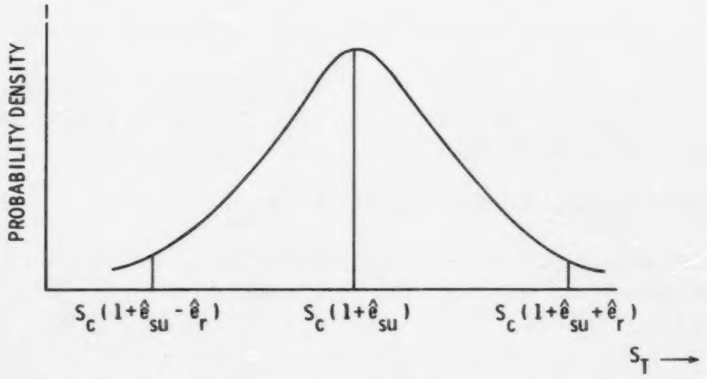

PROBABILITY DENSITY
$S_c(1+\hat{e}_{su}-\hat{e}_r)$
$S_c(1+\hat{e}_{su})$
$S_c(1+\hat{e}_{su}+\hat{e}_r)$
$S_T \longrightarrow$

Figure F-4. Schematic representation of fiducial distribution of S_T for systematic error equal to e_{su}.

The 0.975 probability point is given by

$$S_C(1 + \hat{e}_{su} + \hat{e}_r) = 1.0 \qquad \text{(F-26)}$$

$$S_C = \frac{1}{1 + 17 + 0.14} = 0.0551 \qquad \text{(F-27)}$$

Similarly, the 0.025 probability point is given by

$$S_C(1 + \hat{e}_{su} - \hat{e}_r) = 1.0 \qquad \text{(F-28)}$$

$$S_C = \frac{1}{1 + 17 - 0.14} = 0.0560 \qquad \text{(F-29)}$$

These three points are shown plotted in Figure F-1. Since these points are nearly colinear on this plot, the remainder of the OC curve is approximated by drawing a straight line between the points. Of course, the OC curve must eventually bend over at the ends and approach the horizontal lines P = 1.0 and P = 0 asymptotically. However, these portions of the curve are of little interest here.

"EFFECTIVE" OC CURVES

The "effective" operating characteristic curve shown in Figure 5 (Section 5) can be constructed using the appropriate branches of the curves in Figure F-1. The upper branch of the curve for systematic error equal to e_{su} and the lower branch of the curve for systematic error equal to $-e_{s\ell}$ are connected by a horizontal line.

Alternatively, Figure 5 can be constructed by working directly with an "effective" probability distribution. The "effective" fiducial distribution of S_T is obtained by treating Interval F-6 as an exact 95% confidence interval for S_T. The resulting distribution is shown schematically in Figure F-5. The area under each branch of the density function is 0.5. The 2.5% point of the distribution is $S_C(1 - \hat{e}_{s\ell} - \hat{e}_r)$, and the 97.5% point is $S_C(1 + \hat{e}_{su} + \hat{e}_r)$.

By following the procedure described in the previous subsections, it can be verified that the density function shown in Figure F-5 generates the "effective" OC curve of Figure 5.

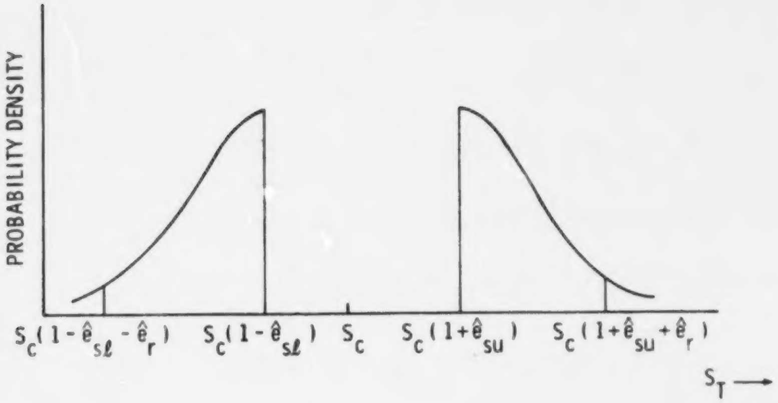


Figure F-5. Schematic representation of "effective" fiducial distribution of S_T.

APPENDIX G

RELATIONSHIP OF SAMPLING AND ANALYSIS BIAS TO SYSTEMATIC ERRORS

In this appendix, the relationship of sampling and analysis bias to systematic errors is developed.

Let β^+ = upper bound on positive bias in sampling and analysis
β^- = upper bound on negative bias in sampling and analysis

where $\beta^+, \beta^- \geq 0$

Then, letting B_{meas} and B_{true} represent the measured and true values of the emission factor,

$$\beta^+ \geq \frac{B_{meas} - B_{true}}{B_{true}} \qquad \beta^- \geq \frac{B_{true} - B_{meas}}{B_{true}}$$

$$B_{true}\left(1 + \beta^+\right) \geq B_{meas} \qquad B_{meas} \geq B_{true}\left(1 - \beta^-\right)$$

$$B_{true} \geq \frac{B_{meas}}{1 + \beta^+} \qquad B_{true} \leq \frac{B_{meas}}{1 - \beta^-}$$

$$\frac{B_{meas}}{1 + \beta^+} \leq B_{true} \leq \frac{B_{meas}}{1 - \beta^-}$$

By definition of $\hat{b}_{su}$ and $\hat{b}_{s\ell}$,

$$B_{meas}\left(1 - \hat{b}_{s\ell}\right) \leq B_{true} \leq B_{meas}\left(1 + \hat{b}_{su}\right)$$

Therefore,

$$\frac{B_{meas}}{1 - \beta^-} = B_{meas}\left(1 + \hat{b}_{su}\right) \qquad \text{(G-1)}$$

and

$$\frac{B_{meas}}{1 + \beta^{+}} = B_{meas}\left(1 - \hat{b}_{s\ell}\right) \tag{G-2}$$

or

$$\frac{1}{1 - \beta^{-}} = 1 + \hat{b}_{su} \tag{G-3}$$

or

$$\frac{1}{1 + \beta^{+}} = 1 - \hat{b}_{s\ell} \tag{G-4}$$

Therefore,

$$\hat{b}_{su} = \frac{\beta^{-}}{1 - \beta^{-}} \tag{G-5}$$

and

$$\hat{b}_{s\ell} = \frac{\beta^{+}}{1 + \beta^{+}} \tag{G-6}$$

APPENDIX H

SOURCE SEVERITY SIMULATION AND PROBABILISTIC SENSITIVITY ANALYSIS

BACKGROUND

In the beginning of the Source Assessment Program, a decision index, "source severity," was developed. For noncriteria pollutants, i.e., pollutants other than particulate matter, sulfur dioxide, nitrogen dioxide, hydrocarbons, and carbon monoxide, stationary source severity was defined as

$$S = \frac{5.5\ Q}{\mathrm{TLV}(h^2)} \qquad \text{(H-1)}$$

where S = stationary source severity
Q = air pollutant emission rate, g/s
TLV = threshold limit value, g/m^3
h = physical emission height, m (typically, chimney or stack height)

This simple and usable expression was derived after a number of conservative, simplifying assumptions had been made. For screening and evaluative purposes, this form of the decision index was deemed appropriate at the time.

During the course of the Source Assessment Program, questions arose regarding the effects of uncertainty in various parameters. Addressed here in one combined analysis are a more complete formulation of the severity model, physical distributions of parameters, and uncertainties in several parameters. In this simulation study are included effects of varying atmospheric stability; variations in wind speed; uncertainties in the ratio of peak concentrations, χ_{peak}, to 24-hr maxima, χ_{24}; uncertainties in the emission factor; uncertainties in lateral dispersion coefficients, σ_y; uncertainties in vertical dispersion coefficients, σ_z; variations in temperature; and resulting variations in plume rise. D. B. Turner's well-documented air dispersion model (5) has been used. The analytical effort to rigorously calculate the effects of all uncertainties is formidable if not impossible. A Monte Carlo simulation was used to investigate effects of uncertainties in input parameters.

DESCRIPTION OF REPRESENTATIVE PLANT

The simulated representative plant was a coal-fired, steam electric utility assumed to be emitting beryllium from one virtual point or stack. Simulated plant constants and fuel parameters are shown in Table H-1. Meteorological data for atmospheric stability and wind speed correspond to central Alabama and were used because they were readily available in a form suitable for this simulation. The meteorological data were the joint frequency of occurrence of atmospheric stability and wind speed. Frequency-distributed variables and their parameters are shown in Table H-2. Since there were only six wind speed ranges and six stability classes, these parameters were not distributed but were used by means of table lookup procedures.

Equations that were used in the severity simulation, ordered in calculational sequence, may be found in Example 1 at the end of this appendix. The source severity calculation for this particular source, emitting beryllium, is shown in Example 2. This severity, subsequently referred to as deterministic severity, was calculated to be 18.6. If plume rise is included in the calculation, the resulting deterministic severity is 3.3. The latter calculation is given in Example 3.

SIMULATION RESULTS

The ground level concentration of a pollutant emitted by an elevated point source is a function of several variables, one of which is the downwind distance. For this initial analysis, a distance was used where the mean of the ground level concentration distribution is a maximum. This distance could not be calculated directly and was approximated by a binary search. It was found to lie between 2 km and 4 km. A distance of 3 km was thus used for all simulations. Results of the baseline simulation, where all variables were distributed, and of probabilistic sensitivity analyses are presented in Table H-3.

GENERAL COMMENTS ON RESULTS

In a Monte Carlo simulation, since distributions are used as inputs and not single deterministic values, the result is also a distribution of values--in this case, a distribution of values for severity. When a probabilistic sensitivity analysis was performed, the variables were fixed one at a time at either a high level or low level. However, since the remaining variables are still distributions of values, the resulting severity will again be a distribution of values.

The sensitivity analyses were performed by fixing each of the variables that were normally distributed at the $+2\sigma$ level and the -2σ level. Thus, in the second row of Table G-3, the notation $S_{\overline{P}0.05}$ designates that the value of P (the ratio of peak to 24-hr

TABLE H-1. PLANT AND FUEL PARAMETERS USED IN SIMULATION

Parameter	Value
Coal consumption rate, g/s	34,532[a]
Stack height, m	100[b]
Barometric pressure, m[b]	1,013
Stack gas temperature, °K	408[b]
Fraction of beryllium in coal, ppm	2.3[b]
Stack I.D., m	4.23[b]
Stack gas velocity, m/s	14.1[b]
TLV of beryllium, g/m^3	2×10^{-6}[c]

[a] 1,092,000 metric tons/yr (nominal value for a 500-MW powerplant). (1 metric ton equals 10^6 grams; conversion factors and metric system prefixes are presented at the end of this report).

[b] Nominal value for a 500-MW powerplant.

[c] Reference 4.

TABLE H-2. ASSUMED DISTRIBUTION PARAMETERS

Variable	Assumed distribution	Mean	Standard deviation	Lower limit	Upper limit
$\frac{X_{peak}}{X_{24\text{-}hr}} = P$	Normal.	6.2[a]	0.93[a]	4.3	8.1
Emission factor	Normal.	0.85	0.347	0.17	1.53
Ambient temperature, °K	Normal.	286.1	4.9	276.5	295.7
σ_z	Normal.	1.0	0.255	0.5	1.5
σ_y	Normal.	1.0	0.051	0.9	1.1
Atmospheric stability	None (table lookup).				
Wind speed	None (table lookup).				

NOTE.—Blanks indicate numerical value is not applicable.

[a] Results derived in Reference 32 from fixed-site ambient sampling data.

---

(32) Montgomery, T. L., and J. H. Coleman. Empirical Relationships Between Time-Averaged SO_2 Concentrations. Environmental Science and Technology, 9(10):953-957, 1975.

TABLE H-3. SIMULATION RESULTS OF PROBABILISTIC SENSITIVITY ANALYSES

	Mean $\overline{S}_x$	Standard deviation	$S_{0.95}$	$P(S \leq 0.1)$	$P(S \leq 1.0)$	$P(S > 1.0)$
Baseline S_0	0.0828	0.788	1.85	0.585	0.841	0.159
$S_{P_{0.05}}$ $P_{0.05} = 4.3$	0.122	1.11	2.70	0.561	0.771	0.229
$S_{P_{0.95}}$ $P_{0.95} = 8.1$	0.065	0.592	1.44	0.597	0.887	0.113
$S_{EF_{0.05}}$ $EF_{0.05} = 0.17$	0.0163	0.153	0.372	0.688	1.00	0.0
$S_{EF_{0.95}}$ $EF_{0.95} = 1.53$	0.148	1.38	3.41	0.541	0.703	0.297
$S_{Ta_{0.05}}$ $Ta_{0.05} = 276.5$	0.076	0.749	1.89	0.586	0.853	0.147
$S_{Ta_{0.95}}$ $Ta_{0.95} = 295.7$	0.086	0.832	2.05	0.568	0.827	0.173
$S_{\sigma_{z_{0.05}}}$ $\sigma_z = \sigma_z - 0.5\ \sigma_z$	0.027	0.367	0.97	0.806	0.954	0.046
$S_{\sigma_{z_{0.95}}}$ $\sigma_z = \sigma_z + 0.5\ \sigma_z$	0.139	1.06	2.28	0.451	0.675	0.325
$S_{\sigma_{y_{0.05}}}$ $\sigma_y = \sigma_y - 0.1\ \sigma_y$	0.097	0.842	2.05	0.576	0.823	0.177
$S_{\sigma_{y_{0.95}}}$ $\sigma_y = \sigma_y + 0.1\ \sigma_y$	0.08	0.689	1.70	0.586	0.859	0.141
Stability Class A	0.0242	0.168	0.34	0.513	1.00	0.0
Stability Class F	9.3×10^{-8}	4.75×10^{-5}	<0.01	1.0	1.0	0.0
Stability D $u_{0.05} = 2.05$	0.0153	0.154	0.65	0.764	0.999	0.001
Stability D $u_{0.95} = 6.99$	0.112	0.72	1.48	0.147	0.832	0.168
$S_{P_{0.5}}$ $P_{0.5} = 6.2$	0.085	0.773	1.62	0.581	0.838	0.162
$S_{EF_{0.5}}$ $EF_{0.5} = 0.85$	0.812	0.776	1.91	0.565	0.824	0.176

maximum concentrations) was fixed at the -2σ level where P is equal to 4.3. The third row shows that for this sensitivity analysis, P was fixed at the $+2\sigma$ level (or 8.1). Since results of the probabilistic simulation are also distributions, several parameters are available for making comparisons. One of the obvious parameters is the mean of the resulting distributions; other parameters are listed in Table H-3 as column headings and include the standard deviation, the severity at the 95% cumulative level, and the probability that the severity is ≤ 0.1, ≤ 1.0, and >1.

All of the resulting distributions are nonnormal. Looking at the equations in Example 1, the mean severity values shift in the appropriate direction. Thus, when the emission factor is low, the mean severity is low; and when the emission factor is high, the mean severity is high with respect to the baseline case.

When the ambient temperature is low, the buoyancy factor in the plume rise equation (Equation H-4) is high; therefore, the plume rise is higher and thus the severity is less than the baseline case. The opposite is true when the ambient temperature is high.

The effect of σ_z is more complex, since σ_z is both in the exponential term as well as in the denominator of the premultiplying factor in the Gaussian plume equation (Equation H-7). The effect of the exponential term overshadows the premultiplier term, as can be seen from the lower value of σ_z yielding a lower value of severity when compared to the baseline case. The opposite is true for higher values of σ_z.

In the case of σ_y, the effect is straightforward since σ_y is used in only the denominator of the premultiplying factor and thus the low values of σ_y give higher values of severity compared to the baseline case. The opposite is true for higher values of σ_y.

For stability Class A, a lower value of mean severity is obtained since values of σ_y and σ_z are large in this stability class. The exponential term thus becomes unity, and only the premultiplying term becomes significant. Since σ_y and σ_z are in the denominator, a lower value of mean severity is obtained compared to the baseline case. For stability Class F, values of σ_y and σ_z are very small, smaller than the plume rise by a significant factor. Thus, for example, if H = 300 m and σ_z = 30 m,

$$\exp\left[-0.5\left(\frac{300}{30}\right)^2\right] = 1.9 \times 10^{-22} \qquad \text{(H-2)}$$

and the severity is much lower than the baseline case.

Stability D was fixed to evaluate the effect of fixed wind speed. Using a low value of wind speed results in a higher plume rise; this in turn gives a lower value of severity. The wind speed term in the premultiplier in the Gaussian plume equation is overriden by the plume rise term. The converse is true when the wind speed is higher.

The last two rows of Table H-3 show results when mean values of the parameters were used. When P was set equal to its mean value of 6.2, a mean severity approximately equal to the baseline case was obtained. The same was true when a mean emission factor of 0.85 was used. No subsequent simulations using mean values of the parameters were run.

SPECIFIC RESULTS

One indicator of sensitivity is the value of S at the 95% cumulative level. These results are in Column 3 of Table H-3. If the true emission factor is higher than the value being used, the highest severity at the 95% cumulative level is the value of 3.4. Low values of P and high values of σ_z are also sensitive variables. The uncertainties in σ_z and σ_y are the main reason for saying that the dispersion model is no better than a "factor of 2 or 3." However, in this simulation, separation of the individual effects and the overall result should be a good approximation of the "true severity" for this source.

The column labeled "P(S > 1.0)" is another indicator of sensitivity and is a way of looking at the cumulative severity distribution from a different perspective. In one case, if concern relates to the magnitude of the severity at the 95% probability level, Column 3 would be used for sensitivity. If, on the other hand, interest involved the magnitude of the probability that S > 1.0, Column 6 would be used. In either case, P, σ_z, and the emission factor are the most sensitive variables.

LARGE SYSTEMATIC ERROR IN EMISSION FACTOR

One final question that was addressed involves the effect of a large systematic error in the emission factor. First, plant and fuel parameters are assumed to be known exactly. While this is not strictly true (e.g., stack gas flow rate and temperature may vary), uncertainties are assumed to be small. The emission factor in this analysis is expressed as a fraction of beryllium that is emitted based on the amount of beryllium in the coal. To preserve the uncertainty distribution on the emission factor and to facilitate computation, the fraction of beryllium is set in coal (2.3 ppm) at 5 times the baseline value (11.5 ppm). Results of this simulation are shown in Table H-4.

CONCLUSIONS DERIVED FROM THE SIMULATION ANALYSIS

If this simulation approach is assumed to yield a "truer" picture of the actual severity at this representative plant (the true severity distribution being obtainable only by long-term sampling for 24-hr beryllium maxima in a comprehensive sampling network), the following observations can be made:

TABLE H-4. SIMULATION OF A LARGE SYSTEMATIC ERROR IN THE EMISSION FACTOR

$\bar{S}_x$	SD	$S_{0.95}$	$P(S \leq 0.1)$	$P(S \leq 1.0)$	$P(S > 1.0)$
0.414	3.94	9.5	0.523	0.62	0.38

- The deterministic severity is conservative by a factor of 225 when compared to the mean simulated severity for the baseline case (18.6/0.0828).
- The largest value of S at the 95% probability is 3.41 for a high emission factor when looking only at emission factor precision. This represents a conservative factor of 5.5 compared to the deterministic severity.
- Even if the systematic error in emissions is a factor of 5, the deterministic severity is conservative by a factor of 2, (18.6/9.5).

Thus, one may conclude that the first decision index, the deterministic source severity, is indeed a worst case estimate and can be used as one of several tools to aid the EPA decision maker in the screening process.

Example 1. Equations Used in the Severity Simulation[a]

$$Q = (\text{CONS})(\text{EF})(\%\text{Be}) \qquad \text{(H-3)}$$

where Q = emission rate, g/s
CONS = coal consumption, g/s
EF = emission factor for beryllium, g/kg
%Be = fraction of beryllium in coal

$$\Delta h = \left(\frac{V_s d}{u}\right)1.5 + (2.68 \times 10^{-3}) p_b \left(\frac{T_s - T_a}{T_s}\right) d \qquad \text{(H-4)}$$

where Δh = plume rise, m
V_s = stack gas exit velocity, m/s
d = inside stack diameter, m
u = wind speed, m/s
p_b = barometric pressure, mb
T_s = stack gas temperature, °K

---

[a] Ordered in calculational sequence.

T_a = air temperature, °K
2.8×10^{-3} = constant with units of $(\mathrm{mb})^{-1}\ \mathrm{m}^{-1}$

$$\sigma_y = Ax^{0.9031} \tag{H-5}$$

where σ_y = lateral dispersion coefficient, m
x = downwind distance, m
A = function of stability class as shown below

Stability class	A
A	0.3658
B	0.2751
C	0.2089
D	0.1471
E	0.1046
F	0.0722

$$\sigma_z = A_1 x^{B_1} + C_1 \tag{H-6}$$

where σ_z = horizontal dispersion coefficient, m
A_1, B_1, and C_1 (for $x > 1{,}000$ m) = functions of stability as shown below

Stability class	A_1	B_1	C_1
A	0.00024	2.094	-9.6
B	0.055	1.098	2.0
C	0.113	0.911	0.0
D	1.26	0.516	-13
E	6.73	0.305	-34
F	18.05	0.18	-48.6

$$\chi = \frac{Q}{\pi \sigma_y \sigma_z u} \exp\left[-0.5\left(\frac{H}{\sigma_z}\right)^2\right] \tag{H-7}$$

where χ = short-term peak concentration, g/m^3
$H = h + \Delta h$ = effective emission height, m
h = physical stack height, m

$$\chi_{24} = \frac{\chi}{P} \tag{H-8}$$

where χ_{24} = maximum 24-hr concentration, g/m^3
P = empirically derived peak to 24-hr maximum ratio

$$S = \frac{\chi_{24}}{(\mathrm{TLV})(8/24)(1/100)} \tag{H-9}$$

where S = source severity
TLV = threshold limit value, g/m^3

Example 2. Source Assessment Deterministic Severity Calculation

$$Q = (\mathrm{CONS})(\mathrm{EF})(\%\mathrm{Be}) \qquad \text{(H-10)}$$

where Q = emission rate, g/s

CONS = coal consumption rate, g/s
= 3.4532×10^4 g/s

EF = emission factor for beryllium
= 0.85

%Be = fraction of beryllium in coal
= 0.23×10^{-5} g Be/g coal

$$Q = (3.4532 \times 10^4)(0.85)(0.23 \times 10^{-5}) = 6.75 \times 10^{-2} \text{ g/s}$$

$$S = \frac{5.5\ Q}{\mathrm{TLV}(h^2)} \qquad \text{(H-11)}$$

$$= \frac{(5.5)(6.75 \times 10^{-2})}{(2 \times 10^{-6})(100)^2} = 18.6$$

Example 3. Deterministic Severity Using Plume Rise Correction

The plume rise is calculated using Equation D-3.[a]

$$\Delta h = (8.892)\frac{Q_H^{0.444}}{u^{0.694}} \qquad \text{(H-3)}$$

where $Q_H \cong 67 \times 10^{-3}\ V_s\ (d^2) P_b \left(\frac{T_s - T_a}{T_s}\right)$ (H-12)

h = 100 m

u = 4.5 m/s

T_s = 407.95 °K

T_a = 286.1 °K

V_s = 14.1 m/s

d = 4.23 m

---

[a] Equation D-1 (uncorrected Holland equation) gives a plume rise of 65 m and a source severity of 6.8. Equation D-4 (corrected Holland equation) gives a plume rise of 195 m and a source severity of 2.1.

Thus,

$$Q_H = (67 \times 10^{-3})(14.1)(4.23)^2(1,013)\left(\frac{408 - 286}{408}\right)$$

$$Q_H = 5,137 \text{ cal/s}$$

$$\Delta h = \frac{8.892(5,137)^{0.444}}{(4.5)^{0.694}}$$

$$\Delta h \cong 139 \text{ m}$$

$$H = h + \Delta h = 239 \text{ m}$$

$$S = \frac{(5.5)(6.75 \times 10^{-2})}{(2 \times 10^{-6})(239)^2}$$

$$S \cong 3.3$$

APPENDIX I

QUANTIFICATION OF UNCERTAINTY IN SOURCE SEVERITY

Inequality 32, derived in Section 5, can be used directly in the decision-making process by computing values of $\hat{e}_\ell$ and $\hat{e}_u$ corresponding to each individual calculated severity, S_C. Alternatively, Inequality 32 can be used to derive general guidelines for decision making in the Source Assessment Program by computing generalized values of $\hat{e}_\ell$ and $\hat{e}_u$. The latter approach is followed in this report. In this appendix, generalized values are derived for each of the individual uncertainties required in the calculation of $\hat{e}_\ell$ and $\hat{e}_u$.

RANDOM UNCERTAINTIES IN AVERAGE CAPACITY AND EMISSION HEIGHT ($\hat{a}_r$ AND $\hat{c}_r$)

For organic and inorganic chemicals source types having small numbers of plants, $\hat{a}_r$ and $\hat{c}_r$ are usually zero since data for all plants are usually available. For source types having large numbers of sources, survey data are generally obtained from a sufficient number of plants to make these terms small. For example, in a source assessment of asphalt hot-mix plants, 600 plants were surveyed from a population of approximately 4,000 plants to yield 95% confidence limits on mean capacity and mean stack height of ±4% and ±3%, respectively, Thus, for most source types, ±5% appears to be a conservative value for these terms. Then

$$\hat{a}_r^{\,2} + 4\,\hat{c}_r^{\,2} \cong 0.01 \qquad \text{(I-1)}$$

SYSTEMATIC UNCERTAINTY IN EMISSION FACTOR ($\hat{b}_{s\ell}$ AND $\hat{b}_{su}$)

Estimates of systematic and random errors involved in sampling and analytical procedures have been determined for a number of standard EPA methods via collaborative testing. The available results are summarized in Table I-1 (33-43). This table shows that average biases for standard EPA methods range rom -35% to +10%.

---

(33) Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for Determination of Stack Gas Velocity and Volumetric Flow Rate in Conjunction with EPA Method 5. EPA-650/4-74-033 (PB 240 342), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 40 pp.

(continued)

TABLE I-1. COLLABORATIVE TEST RESULTS FOR EPA SAMPLING AND ANALYTICAL METHODS

EPA Method	Bias, % of standard concentration	Precision (standard deviation), % of mean concentration		95% (2σ) confidence limits, % of mean[a]		Minimum detectable limit	Reference
		Within lab	Between lab	Within lab	Between lab		
2		5.5	5.6	11	11.2		36
3		10 to 30	15 to 35	20 to 60	30 to 70		37
5		10 to 30	20 to 40	20 to 60	40 to 80		38-41
6	0	4	5.5	8	11	3 ppm	38,42
7	0	7	10	14	20		38,43,44
8	-2 (Analysis only.)	60	65	120	130		38
9	+1.4% Opacity.	2% Opacity.	2.5% Opacity.	4% Opacity.	5% Opacity.		38
10	-2 to +10	2 to 7	4 to 13	4 to 14	8 to 26	20 ppm	38,45
104	-20	44	58	88	116		38,46
High volume.		3	3.7	6	7.4	3 mg	38
6 (Ambient air.)	0	2 (At 250 $\mu g/m^3$) 1.3 (At 1,000 $\mu g/m^3$)	4 (At 250 $\mu g/m^3$) 2.6 (At 1,000 $\mu g/m^3$)	4 (At 250 $\mu g/m^3$) 2.6 (At 1,000 $\mu g/m^3$)	8 5.2	25 $\mu g/m^3$	38
Chemiluminescent (NO_x), ambient air.)	0	7 to 8 (At 100 $\mu g/m^3$)		14 to 16 (At 100 $\mu g/m^3$)		10 $\mu g/m^3$	38
Chemiluminescent (photochemical oxidants, ambient air).	-35 to -15 (From 0.05 ppm to 0.50 ppm.)	9.2 (At 0.05 ppm) 5.9 (At 0.1 ppm) 3.2 (At 0.5 ppm)		18 (At 0.1 ppm) 12 (At 0.1 ppm) 6.5 (At 0.5 ppm)		0.3 mg/m^3	38
NDIR (ambient air).	+2.5	20 (At 3 mg/m^3)	30 (At 3 mg/m^3)	40 (At 3 mg/m^3)	60 (At 3 mg/m^3)	0.3 mg/m^3	38

NOTE.—Blanks indicate no information available.

[a] Values listed are for a single measurement. For n measurements, the values should be divided by $\sqrt{n}$.

Biases were determined by measuring a gas of known composition (standard) under controlled conditions. Values reported are averages of a number of runs by various collaborators at different levels of pollutant concentration. Since bias can vary greatly among collaborators and with pollutant concentration, the range of the individual measurements would be more appropriate for determining systematic error bounds. For example, in the measurement of carbon monoxide by EPA Method 10 (44), bias in individual tests ranged from -12% to +18%, and there was a definite trend from positive to negative bias with increasing CO concentration. Hence, more realistic error bounds would be obtained by using the limits -12% and +18% for bias rather than the average values of -2% to +10%. For the purpose of the present analysis, however, the average values will be used.[a]

---

[a] The value of -35% for negative bias is probably a conservative estimate for source assessment anyway, since it corresponds to a method for sampling ambient air rather than a stack sampling method.

(continued)

(34) Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for Stack Gas Analysis and Determination of Moisture Fraction with Use of Method 5. EPA-650/4-73-026 (PB 236 929), U.S. Environmental Protection Agency, Research Triangle Park, North Carolnia, 1974.

(35) Smith, F., and J. Buchanan. IERL-RTP Data Quality Manual. EPA-600/2-76-159, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1976.

(36) Hamil, H. F., and D. E. Camann. Collaborative Study of Method for the Determination of Particulate Matter Emissions from Stationry Sources (Portland Cement Plants). EPA-650/4-74-029 (PB 237 346), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 54 pp.

(37) Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for the Determination of Particulate Matter Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators). EPA-650/4-74-021 (PB 234 150), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 36 pp.

(38) Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for the Determination of Particulate Matter Emissions from Stationary Sources (Municipal Incinerators). EPA-650/4-74-022 (PB 234 151), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 37 pp.

(39) Hamil, H. F., and D. E. Camann. Collaborative Study of Method for the Determination of Sulfur Dioxide Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators). EPA-650/4-74-024 (PB 238 293), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1973. 64 pp.

(continued)

From Appendix F,

$$\hat{b}_{su} = \frac{\beta^-}{1 - \beta^-} \qquad \text{(I-2)}$$

$$\hat{b}_{s\ell} = \frac{\beta^+}{1 + \beta^+} \qquad \text{(I-3)}$$

where β^- and β^+ are bounds on the negative and positive bias, respectively. Substituting $\beta^- = 0.35$ and $\beta^+ = 0.10$ yields

$$\hat{b}_{su} \cong 0.5$$

$$\hat{b}_{s_\ell} \cong 0.1$$

RANDOM UNCERTAINTY IN EMISSION FACTOR ($\hat{b}_r$)

The random uncertainty in emission factor will be treated as an independent variable since it can (theoretically) be controlled by specifying the number of samples to be collected and analyzed.

(continued)

(40) Hamil, H. F., and D. E. Camann. Collaborative Study of Method for the Determination of Nitrogen Oxide Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators). EPA-650/4-74-025 (PB 238 555), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1973. 102 pp.

(41) Hamil, H. F., and R. E. Thomas. Collaborative Study of Method for the Determination of Nitrogen Oxide Emissions from Stationary Sources (Nitric Acid Plants). EPA-650/4-74-028 (PB 236 930), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 41 pp.

(42) Constant, P. C., G. Scheil, and M. C. Sharp. Collaborative Study of Method 10--Reference Method for Determination of Carbon Monoxide Emissions from Stationary Sources--Report of Testing. EPA-650/4-75-001 (PB 241 284), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1975.

(43) Constant, P. C., and M. C. Sharp. Collaborative Study of Method 104--Reference Method for Determination of Beryllium Emissions from Stationary Sources. EPA-650/4-74-023 (PB 245 011), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1974. 94 pp.

(44) Title 40--Protection of Environment. Chapter 1--Environmental Protection Agency--Part 60--Standards of Performance for New Stationary Sources. Subchapter C--Air Programs. Method 10--Determination of Carbon Monoxide Emissions from Stationary Sources. Federal Register, 39:9319-9321, 1974.

The range over which $\hat{b}_r$ can be expected to vary can be determined by reference to Table I-1, which lists the precision of standard EPA methods.

The precision of each measurement method was determined by an analysis of variance (ANOVA) of collaborative test results obtained under field conditions. The "within-laboratory" standard deviation is obtained from the residual, or error, variance in the standard ANOVA procedure. It is an estimate of the imprecision inherent in the sampling and analytical procedures and is the imprecision which would be obtained if repeat measurements were performed under identical conditions by the same personnel using the same equipment. The "between-laboratory" standard deviation is obtained from the total variability in the test data, and is a measure of the total imprecision due to 1) the inherent imprecision in the method, 2) differences in personnel and equipment among laboratories, and 3) any process variations which occurred during the tests. In the absence of process variation, the difference of the between-laboratory and within-laboratory variances is a measure of the variability due to differences in factors such as personnel or equipment and is designated "laboratory bias." (Note that laboratory bias is a component of imprecision rather than bias.) In this case,

$$\sigma^2_{\text{between labs}} = \sigma^2_{\text{within lab}} + \sigma^2_{\text{laboratory bias}} \qquad \text{(I-4)}$$

Assuming that errors are normally distributed and that estimates of standard deviations are based on a large number of degrees of freedom (The latter assumption is not justified in some cases due either to poor experimental design or to a limited number of usable data points.), the 95% confidence limits are approximately equal to twice the standard deviation. These values are listed in Table I-1 for both the within-laboratory and between-laboratory components.

Usually, at least three samples are collected and analyzed. Thus, dividing the 95% confidence limits by $\sqrt{3}$, the largest random uncertainty in emission factor should be about 75%. Larger values could result from large process variations, however. Measurement of trace elements and polycyclic organic material (POM) may also result in larger uncertainties due to the relatively large imprecision of analytical methods used for these species. For the purpose of this analysis, it will be assumed that the random uncertainty in emission factor varies between 0 and 100%; i.e.,

$$0 \le \hat{b}_r \le 1.0 \qquad \text{(I-5)}$$

SYSTEMATIC UNCERTAINTY IN EFFECTIVE EMISSION HEIGHT ($\hat{c}_{s\ell}$ AND $\hat{c}_{su}$)

Bounds on the systematic error in effective emission height due to use of semiempirical equations to estimate plume rise are derived in Appendix C. The bounds corresponding to the modified Holland equation (Equation C-4) are used in the present analysis:

$$\hat{c}_{su} \cong 0.1$$

$$\hat{c}_{s\ell} \cong 0.5$$

UNCERTAINTY IN CORRECTION FACTOR FOR DISPERSION EQUATION (m)

There are no experimental data which can be compared with dispersion calculations under the conditions of interest in this report; i.e., fixed-source parameters, wind speed of 4.5 m/s, and Class C stability. However, recent measurements by Guzewich and Pringle (45) permit such a comparison for (approximately) fixed-source parameters and Class C stability. A total of 33 measurements was made under Class C stability conditions with wind speeds ranging from 1.6 m/s to 13.2 m/s. Emission rates varied from 8.4 g/s to 9.8 g/s;[a] other source parameters remained constant during the measurements.

The distribution of the ratio of measured to calculated concentration is shown in Figure I-1.[b] Calculated values were obtained using simple Gaussian dispersion theory together with Briggs' plume rise formula (26). Data are approximately log-normally distributed with a mean of 1.0 and a variance of 0.3136. The 1% and 99% points are approximately equal to 1/3 and 3, respectively.

Data variability can be assigned to variability in the following factors:

- Measured values
- Wind speed
- Emission rate
- Estimated plume rise
- Other atmospheric parameters (other than wind speed and stability class)

Only the last of these contributions is pertinent to the present analysis. That is, error bounds for the correction factor, γ, based on Figure I-1, should be on the conservative side.

---

[a] Emission rates were accurately known since the "pollutant" was a tracer substance injected into the stack gas stream.

[b] Raw data were supplied by R. L. Hanson, Office of the Project Manager for Chemical Demilitarization and Installation Restoration, U.S. Department of the Army, Aberdeen Proving Ground, Maryland.

---

(45) Guzewich, D. C., and W. J. B. Pringle. Validation of the EPA-PTMTP Short-Term Gaussian Dispersion Model. Journal of the Air Pollution Control Association, 27(6):540-542, 1977.

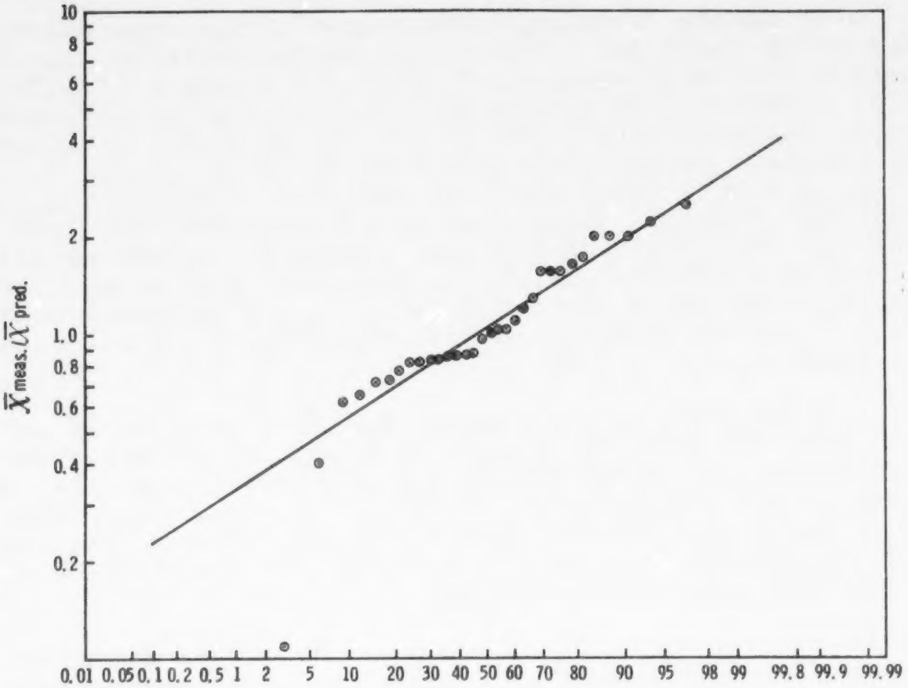


Figure I-1. Comparison of measured and predicted values of ground level concentration.

Therefore, the value m = 3 will be used in the analysis as a conservative estimate.

The above "factor of three" is consistent with the best case estimate given by Turner (5). This factor corresponds to the estimation of the pollutant concentration at a given point downwind of the source rather than estimation of the maximum concentration, $\overline{X}$max. Theoretical and empirical analyses indicate that a "factor of two" may be more appropriate for estimation of $\overline{X}$max (46-47). However, the "factor of three" will be used here in order to take into account uncertainty in the averaging time correction (see following subsection).

Data in Figure I-1 indicate that the uncertainty in the dispersion equation could be eliminated by appropriately modifying the

---

(46) Scriven, R. A. Variability and Upper Bounds for Maximum Ground Level Concentrations. Philosophical Transactions of the Royal Society of London, Series A, 265:209-220, 1969.

(47) Weber, A. H. Atmospheric Dispersion Parameters in Gaussian Plume Modling, Part I, Review of Current Systems and Possible Future Developments. EPA-600/4-76-030a (PB 257 893), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, July 1976. 58 pp.

definition of source severity. The value of $\bar{\chi}_{max}$ was defined as the value which would result from the source emitting into an atmosphere with a wind speed of 4.5 m/s and Class C stability. If a large number of tests were made holding source parameters, wind speed, and stability class fixed, but allowing other atmospheric parameters to vary,[a] a distribution of values of $\bar{\chi}_{max}$ rather than a single value would be obtained. Figure I-1 indicates that the Gaussian dispersion model correctly predicts the median of this distribution (median correction factor equals 1.0). Therefore, by specifying the value of $\bar{\chi}_{max}$ to be used in the definition of source severity as the above ensemble median, inherent uncertainty associated with the dispersion model would be eliminated.

Of course, in interpreting the above modified version of source severity, it would have to be kept in mind that actual ground level concentrations under conditions of Class C stability and wind speed of 4.5 m/s could be "a factor of three" higher or lower than the median value. Hence, "actual" severities could be higher or lower than the nominal value by a "factor of three" under these atmospheric conditions.

From the standpoint of decision making, the modified version of source severity would result in a decision based on the median value of $\bar{\chi}_{max}$ for Class C stability and wind speed of 4.5 m/s rather than on the upper 97.5% value of $\bar{\chi}_{max}$.

UNCERTAINTY IN AVERAGING TIME CORRECTION FACTOR $(t_o/t)^{0.17}$

The averaging time correction factor given by Turner (5), $(t_o/t)^{0.17}$, is a semiempirical relationship derived from data on lateral and vertical diffusion coefficients in steady winds. It applies to situations in which the mean wind direction remains constant over the period of interest. Thus, it corresponds to a worst case situation with respect to pollutant concentration.

A number of studies have been conducted to determine the relationship between measured concentration and averaging time for ambient air monitors located at fixed sites (29, 32, 48). Results of these studies tend to predict considerably lower mean concentrations than does the above correction factor, which is to be expected since the mean wijd direction does not, in general, remain constant over an extended period of time. The diversity

---

[a] For example, stability Class C defines a range of atmospheric stability conditions rather than a single unique condition. Previously, it was assumed that other atmospheric parameters were held fixed so that a unique atmospheric state (and a corresponding unique value of $\bar{\chi}_{max}$) was obtained.

---

(48) Shiruaikar, V. V., and P. R. Patel. Long Term Statistics of Peak/Mean Concentrations From A Point Source. Atmospheric Environment, 11(4):387-389, 1977.

of results obtained in the three studies cited above also emphasizes the fact that the relationship between concentration and sampling time is strongly dependent upon local meteorological conditions; e.g., wind direction statistics.

For the purpose of the present analysis, it is assumed that the mean wind direction remains constant during the period of interest. This assumption is consistent with the concept of a fixed-state receiving atmosphere assumed previously. The uncertainty associated with the averaging time correction factor under these conditions is then required. In lieu of other information, this uncertainty is assumed to be included in the uncertainty associated with the dispersion equation, and the latter uncertainty is conservatively taken to be a "factor of three" (see preceding subsection).

The effect on source severity of relaxing the assumption of constant mean wind direction is considered in Appendix G.

UNCERTAINTY IN ACCEPTABLE CONCENTRATION ($\hat{d}_{su}$ AND $\hat{d}_{s\ell}$)

It is assumed that for each noncriteria pollutant, i, the "true" value, F_T, and calculated value, F_C, of the "acceptable" concentration are given by

$$\left(F_T\right)_i = (\mathrm{TLV})_i\, G_i \qquad \text{(I-6)}$$

$$\left(F_C\right)_i = (\mathrm{TLV})_i\, \overline{G} \qquad \text{(I-7)}$$

where $(\mathrm{TLV})_i$ = TLV of pollutant i as listed in Reference 4

G_i = conversion factor which converts $(\mathrm{TLV})_i$ to an equivalent PAAQS

$\overline{G}$ = an "average" conversion factor

Since $(\mathrm{TLV})_i$ is taken as the nominal value listed in Reference 4, there is no uncertainty associated with this value in the context of the present analysis. All variability due to the methodology used in establishing TLV's is associated with the conversion factors, G_i.

From Equations I-6 and I-7 are obtained

$$\frac{\left(F_T\right)_i}{\left(F_C\right)_i} = \frac{G_i}{\overline{G}}$$

$$\ln \frac{\left(F_T\right)_i}{\left(F_C\right)_i} = \ln G_i - \ln \overline{G} \qquad \text{(I-8)}$$

It is assumed that the ln G_i are normally distributed (i.e., the G_i are log-normally distributed) over the population of pollutants with mean μ and variance σ^2. It then follows from Equation I-8 that $\ln(F_T/F_C)$ is normally distributed with mean $\mu - \ln \overline{G}$ and variance σ^2. If it is assumed that $\overline{G}$ is the geometric mean of the distribution of G_i, then $\ln \overline{G} = \mu$, and the distribution of $\ln(F_T/F_C)$ has a mean of zero.

Now it is assumed that criteria pollutants constitute a random sample of size four from the population of all pollutants. From Table E-1, the values shown in Table I-2 are obtained.

TABLE I-2. CONVERSION FACTORS FOR CRITERIA POLLUTANTS

Pollutant	G_i	$\ln G_i$
Particulate matter	0.0260	-3.650
Sulfur oxides (SO_x)	0.0281	-3.573
Carbon monoxide (CO)	0.158	-1.845
Nitrogen oxides (NO_x)	0.0411	-3.191

Using these values, the estimated mean and standard deviation of the ln G_i are

$$\bar{x} = -3.064$$

$$s = 0.8395$$

Thus, the estimate for $\overline{G}$ is

$$\overline{G} = e^{\mu} \cong e^{\bar{x}} = 0.0467 \qquad \text{(I-9)}$$

The estimated 95% confidence interval for $\ln(F_T/F_C)$ is ±1.96 s; i.e.,

$$-1.645 \leq \ln\left(\frac{F_T}{F_C}\right) \leq 1.645$$

$$0.193 \leq \frac{F_T}{F_C} \leq 5.18$$

$$0.193 \, F_C \leq F_T \leq 5.18 \, F_C \qquad \text{(I-10)}$$

By definition of $\hat{d}_{su}$ and $\hat{d}_{s\ell}$,

$$F_C(1 - \hat{d}_{s\ell}) \leq F_T \leq F_C(1 + \hat{d}_{su}) \qquad \text{(I-11)}$$

From Equations I-10 and I-11,

$$\hat{d}_{su} \cong 4.2$$

$$\hat{d}_{s\ell} \cong 0.81$$

In the Source Assessment Program, $\overline{G} = 0.0033$, so that $\ln \overline{G} = -5.70$. In this case, the estimated 95% confidence interval for F_T/F_C is given by

$$[-3.064 - (-5.70)] \pm 1.96\ s$$

That is,

$$2.64 - 1.645 \leq \ln \frac{F_T}{F_C} \leq 2.64 + 1.645$$

$$2.70 \leq \frac{F_T}{F_C} \leq 72.6$$

$$2.70\ F_C \leq F_T \leq 72.6\ F_C \qquad \text{(I-12)}$$

Comparing Equations I-11 and I-12 yields[a]

$$\hat{d}_{su} = 71.6$$

$$\hat{d}_{s\ell} = 0$$

Thus, there are three sets of values for the uncertainty in "acceptable" concentration.

Case A. No uncertainty in acceptable concentration.

$$\hat{d}_{s\ell} = \hat{d}_{su} = 0$$

Case B_1. Acceptable concentration uncertain; TLV conversion factor equal to geometric mean value.

$$\hat{d}_{su} = 4.2$$

$$\hat{d}_{s\ell} = 0.81$$

Case B_2. Acceptable concentration uncertain; TLV conversion factor equal to 1/300.

---

[a] Note that $\hat{d}_{s\ell} \geq 0$ by definition.

$\hat{d}_{su} = 71.6$

$\hat{d}_{s\ell} = 0$

APPENDIX J

EXAMPLE CALCULATIONS FOR SOURCE ASSESSMENT: CARBON BLACK MANUFACTURE

In this appendix, the principles developed in the body of the report are applied to a source assessment study of carbon black manufacturing operations (49). In this study, emission factors for the main process vent were measured at a typical carbon black plant. Mean emission factors and 95% confidence limits for the mean values (computed from the experimental data) are listed in Table J-1. These values were taken directly from Reference 52. Also listed in Table J-1 are the random and systematic uncertainties associated with the measured emission factors. The values of $\hat{b}_r$ are simply the 95% confidence limits for the mean emission factors. The systematic uncertainties for nitrogen oxides and carbon monoxide were obtained by applying Equations G-5 and G-6 to the biases listed in Table I-1 for EPA Methods 7 and 10, respectively. The systematic uncertainties for the other species in Table J-1 were equated with the generalized values given in Table 3 (Section 5), since no information is available concerning biases in the methods used to measure emissions of these compounds.

TREATMENT OF S_T AS A FIXED PARAMETER

The random uncertainties in average plant capacity and average emission height ($\hat{a}_r$ and $\hat{c}_r$) were set equal to zero since data were available for all plants in the industry. Values given in Table 3 (Section 5) were used for the uncertainties in plume rise, dispersion equation, and "acceptable" concentration.

For each pollutant listed in Table J-1, the following quantities were computed:

- Estimated severity, S_C (including plume rise).
- Lower and upper bounds for the true severity, S_T.
- The critical test value, S_C^*.

---

(49) Serth, R. W., and T. W. Hughes. Source Assessment: Carbon Black Manufacture. EPA-600/2-77-107k, U. S. Environmental Protection Agency, Research Triangle Park, North Carolina, October 1977. 244 pp.

Results for criteria pollutants are presented in Table J-2. Criteria pollutants correspond to Case A of Section 5; i.e., no uncertainty in "acceptable" concentration. The critical test value derived in Section 5 by the generalized analysis was 0.05. The values for nitrogen oxides and carbon monoxide are significantly higher than the generalized value because of the smaller systematic uncertainties in emission factors for these two species.

TABLE J-1. EMISSION FACTORS AND ASSOCIATED UNCERTAINTIES FOR CARBON BLACK PROCESS VENT

Material emitted	Measured emission factor, g/kg	$\hat{b}_r$	$\hat{b}_{s\ell}$	$\hat{b}_{su}$
Particulate matter	0.11 ± 70%	0.70	0.1	0.5
Nitrogen oxides	0.28 ± 15%	0.15	0	0
Hydrocarbons	50 ± 48%	0.48	0.1	0.5
Carbon monoxide	1,400 ± 19%	0.19	0.09	0.02
Hydrogen sulfide	30 ± 82%	0.82	0.1	0.5
Carbon disulfide	30 ± 76%	0.76	0.1	0.5
Carbonyl sulfide	10 ± 99%	0.99	0.1	0.5
Isobutane	0.1 ± 80%	0.80	0.1	0.5
n-Butane	0.27 ± 57%	0.57	0.1	0.5
POM (total)	0.002 ± 52%	0.52	0.1	0.5

TABLE J-2. EXAMPLE CALCULATIONS FOR CRITERIA POLLUTANTS (CASE A: NO UNCERTAINTY IN ACCEPTABLE CONCENTRATION)

Material emitted	Calculated severity, S_C	Uncertainty interval $S_C - e_\ell \le S_T \le S_C + e_u$	Critical test value, S_C^*
Particulate matter	0.004	$0 \le S_T \le 0.075$	0.054
Nitrogen oxides	0.032	$0.004 \le S_T \le 0.39$	0.082
Hydrocarbons	4.2	$0 \le S_T \le 78$	0.054
Carbon monoxide	0.36	$0.022 \le S_T \le 4.5$	0.080

Comparing the values of S_C and S_C^* in Table J-2, it would be concluded that the source type is a potential candidate for emissions reduction with regard to emissions of hydrocarbons and carbon monoxide (since $S_C > S_C^*$ for these species), but not with regard to nitrogen oxides and particulate emissions (since $S_C < S_C^*$ for these species).

For noncriteria pollutants, calculations were made for both Case B_1 (TLV conversion factor of 0.0467) and Case B_2 (TLV conversion factor of 0.0033). Results of these calculations are given in Tables J-3 and J-4, respectively. The critical test values obtained by the generalized analysis in Section 5 were 0.01 for Case B_1 and 0.05 for Case B_2. The critical test values listed in Tables J-3 and J-4 are essentially the same as the generalized results since the same uncertainty values were used in both cases.

TABLE J-3. EXAMPLE CALCULATION FOR NONCRITERIA POLLUTANTS (CASE B_1: TLV CONVERSION FACTOR = 0.0467)

Material emitted	Calculated severity, S_C	Uncertainty interval $S_C - e_\ell \le S_T \le S_C + e_u$	Critical test value, S_C^*
Hydrogen sulfide	0.28	$0 \le S_T \le 27$	0.0105
Carbon disulfide	0.11	$0 \le S_T \le 11$	0.0105
Carbonyl sulfide	0.030	$0 \le S_T \le 2.9$	0.0105
Isobutane	0.000014	$0 \le S_T \le 0.0013$	0.0105
n-Butane	0.000042	$0 \le S_T \le 0.0040$	0.0105
POM (total)	0.030	$0 \le S_T \le 2.9$	0.0105

TABLE J-4. EXAMPLE CALCULATION FOR NONCRITERIA POLLUTANTS (CASE B_2: TLV CONVERSION FACTOR = 0.0033)

Material emitted	Calculated severity, S_C	Uncertainty interval $S_C - e_\ell \le S_T \le S_C + e_u$	Critical test value, S_C^*
Hydrogen sulfide	4.0	$0 \le S_T \le 75$	0.053
Carbon disulfide	1.5	$0 \le S_T \le 28$	0.053
Carbonyl sulfide	0.42	$0 \le S_T \le 8.0$	0.053
Isobutane	0.00020	$0 \le S_T \le 0.0038$	0.053
n-Butane	0.00060	$0 \le S_T \le 0.011$	0.054
POM (total)	0.42	$0 \le S_T \le 7.8$	0.054

Comparing the values of S_C and S_C^* in Table J-3, it would be concluded that the source type is a potential candidate for emissions reduction with regard to emissions of hydrogen sulfide, carbon disulfide, carbonyl sylfide, and POM, but not with regard to emissions of isobutane or n-butane. The same conclusions would be reached using the values in Table J-4.

GLOSSARY

accuracy: Closeness between true and measured values which is characteristic of a given measurement process.

atmospheric stability class: Class used to designate degree of turbulent mixing in the atmosphere.

bias: Magnitude and direction of the tendency of a given measurement process to measure a quantity other than the nominal or intended quantity.

criteria pollutant: Emission species for which an ambient air quality standard has been established.

decision index: Numerical index used as an aid in decision making.

emission factor: Weight of material emitted to the atmosphere per unit weight of product produced.

expected value: First moment about the origin (i.e., the mean) of the probability distribution of a random variable.

Monte Carlo simulation: Simulation in which variables are represented by probability distributions.

national emissions burden: Decision index defined as the ratio of emissions of a criteria pollutant from a given source type to emissions of the same pollutant from all stationary sources nationwide.

Neyman-Pearson theory: Most widely accepted theory of statistical inference in which population parameters are regarded as fixed.

noncriteria pollutant: Emission species for which no ambient air quality standard has been established.

operating characteristic curve: Curve which gives the (fiducial) probability that the true severity is less than one as a function of the calculated severity.

plume rise: Distance above point of emission where plume rises due to its momentum and buoyancy.

precision: Closeness together, or lack of scatter, in replicate measurements which is characteristic of a given measurement process.

random error: Component of the total error which results from imprecision of the measurement process.

random uncertainty: Component of the total uncertainty which results from random error in the measured value.

reasonably available control technology (RACT): Available, economically feasible control technology required for attainment of national ambient air quality standards (as specified in Code of Federal Regulations, Title 40, Chapter I, Part 51. Requirements for Preparation, Adoption, and Submittal of Implementation Plans).

representative source: Hypothetical source having typical values of parameters which is used to characterize an industry or source type.

source severity: Decision index defined as the ratio of maximum time-averaged ground level pollutant concentration to an acceptable pollutant concentration.

systematic error: Component of the total error which results from bias in the measurement process.

systematic uncertainty: Component of the total uncertainty which results from the systematic error in the measured value.

Type I error: Error in decision making which results from rejecting the null hypothesis when it is true.

Type II error: Error in decision making which results from accepting the null hypothesis when it is false.

uncertainty: Measure of the lack of knowledge of the true value of a quantity due to imperfect data; i.e., due to errors in the measured values of the quantity.

CONVERSION FACTORS AND METRIC PREFIXES (50)

CONVERSION FACTORS

To convert from	To	Multiply by
Degree Celsius (°C)	Degree Fahrenheit	$t^\circ_F = 1.8\ t^\circ_C + 32$
Degree Kelvin (°K)	Degree Celsius	$t^\circ_C = t^\circ_K - 273.15$
Gram/second (g/s)	Pound/hr	7.937
Joules (J)	Calories	2.388×10^{-1}
Kilogram (kg)	Pound-mass (avoirdupois)	2.205
Kilometer (km)	Mile	6.214×10^{-1}
Meter (m)	Foot	3.281
Meter^2 (m^2)	Foot^2	1.076×10^{1}
Meter^3 (m^3)	Foot^3	3.531×10^{1}
Meter/second (m/s)	Mile/hr	2.237×10^{3}
Metric ton	Pound-mass	2.205×10^{3}
Metric ton	Kilogram	1.000×10^{3}
Metric ton	Ton (short, 2,000 pound mass)	1.102
Newtons (n)	Dyne	1.000×10^{5}
Pascal (Pa)	Millibars	1.000×10^{-2}
Pascal (Pa)	Pound-force/inch^2 (psi)	1.450×10^{-4}
Watts (W)	Calories/min	1.434×10^{1}

METRIC PREFIXES

Prefix	Symbol	Multiplication factor	Example
Kilo	k	10^3	1 kPa = 1 x 10^3 pascals
Mega	M	10^6	1 MJ = 1 x 16^6 joules
Micro	μ	10^{-6}	1 μg = 1 x 10^{-6} gram

---

(50) Standard for Metric Practice. ANSI/ASTM Designation E 380-76$^{\varepsilon}$, IEEE Std 268-1976, American Society for Testing and Materials, Philadelphia, Pennsylvania, February 1976. 37 pp.